Some Branches of Mathematics from the Fuzzy Set Theory Perspective.

José Rubén Niño Quevedo
Matemático

Universidad Nacional de Colombia
Facultad de Ciencias
Departamento de Matemáticas
Bogotá, D.C.
2018

Some Branches of Mathematics from the Fuzzy Set Theory Perspective.

José Rubén Niño Quevedo
Matemático

Trabajo Final de Maestría
para optar el título de
MSc. Matemáticas

Advisor
Rodrigo de Castro Korgi, Ph.D.

Research line
Matemática Difusa

Universidad Nacional de Colombia
Facultad de Ciencias
Departamento de Matemáticas
Bogotá, D.C.

2018

Title in English

Some Branches of Mathematics from the Fuzzy Set Theory Perspective.

Título en español

Algunas Ramas de las Matemáticas desde la perspectiva de la Teoría de los Conjuntos Difusos.

Abstract: In this document some concepts and ideas in classic mathematics are extended by using the concepts and methods of the Fuzzy sets theory; the attention focused on three areas: Logic, Measure theory and Theory of Computation.

Resumen: En este documento, algunos conceptos e ideas en matemáticas clásicas se amplían utilizando los conceptos y métodos de la teoría de conjuntos difusos; la atención se centró en tres áreas: lógica, teoría de la medida y teoría de Computación.

Keywords: Fuzzy Sets, Arithmetic, Logic Measure, Theory of Computation.

Palabras clave: Conjuntos Difusos, Aritmética, Lógica , Medida, Teoría de la Computación.

Acceptation Note

Thesis Work

Jury

Jury

Advisor
Rodrigo de Castro Korgi

Bogotá, D.C., 2018

Dedication.

The present work is for the National University of Colombia, its members and to whom are motivated to give more they can.

Acknowledgement.

The author expresses his thanks to everyone who supported and encouraged him in the preparation of this work, especially to his parents Ana Graciela Quevedo Aldana and José Ángel Niño, his “niña” Liliana Esperanza Niño Quevedo, his director, the professor PhD. Rodrigo de Castro Korgi and, his friends: Carlos Isaac Zainea, Ana María Cruz, Omar Andrés Cuervo and his loved MP.

Also, the author expresses his gratitude to his other classmates and teachers who shared and collaborated until now, specially for the professor PhD. Jonatan Gomez, since their comments helped or contributed to the development or improvement of this work.

Contents

List of Tables

List of Figures

Introduction

Mathematical models are often considered to describe and solve a real-world problem because they are "simple and precise" frameworks that give a "good approximation of reality". In practice, it is common to deal with complexity and uncertainty when we start the task of translating the "real world" phenomena into the language of mathematics. Setting variables or relationships is one of the first steps to model a situation but in several cases, they are subject to change at the point of becoming ambiguous, unpredictable or vague. Reducing this type of uncertainty leads to increase the complexity but, at the same time, this increment makes our ability to be precise to decrease. In order to shorten this uncertainty, words, expressions even subjective judgements are more appropriated than figures, however, lots of these linguistic terms cannot be formally treated by traditional approaches.

In 1965, with the purpose of dealing with phenomena related to human thinking, Lofti Zadeh proposed a mathematical object called fuzzy set [79] which provides a natural relation between linguistic representations and numerical ones. Zadeh said that getting a better understanding implies to explore the use of variables whose values are words or sentences expressed in a natural language instead of figures [81, 85]. Thus, Fuzzy Set Theory (FuSeTh) arose not only as a natural but mathematical way to handle the uncertainty tied to vagueness [21], approximate reasoning [12] and human thinking like common sense, judgements or concepts expressed in natural language. This idea has made FuSeTh an interesting research area since the use of linguistic variables reduces the complexity of the models in a wide variety of cases [12, 25, 45, 95, 96] and even, fuzzy sets can be considered suitable to process data derived from vague sources [74].

Although FuSeTh has never been an invitation to fuzzy thinking [97], people often joke when they hear of words like fuzzy logic, fuzzy algebra, fuzzy analysis. They probably think that it is fuzzy indeed. However, this branch of mathematics is an effort to formalize the way we deal with the uncertainty due to vagueness or concepts expressed in natural languages [21]. In addition, fuzzy sets can be considered as a generalization of the classical notion for the word set since they catch the idea of partial membership but, the idea of partial membership, makes people confuse FuSeTh with the probability theory. It is worth indicating that fuzziness is both conceptually and formally different from the concept of probability. Fuzzy sets are membership functions whereas probability is a set function. Fuzzy sets are used in Possibility theory to say whether an event may occur and with what degree; on the contrary, the probability is about whether an event will occur. Moreover, we highlight that FuSeTh provides a way for handling uncertainty that escapes from the scope of the stochastic models, that it is a suitable tool for handling uncertainty due to imprecision and vagueness instead of randomness and also, that a linguistic variable can be interpreted as a dictionary in the sense it translates linguistic terms, which may be composed of a sentences whose meaning can contain ambiguity and modifiers that change the meaning of predicates, into mathematical objects known as fuzzy sets [6].

This document is thought of as an introduction to three branches of fuzzy mathematics: Fuzzy logic, Fuzzy measure theory and the theory of fuzzy computation. To achieve this objective we read and collected several ideas from different authors in order to present a self-contained work that intends to motivate readers to explore and research in this new branch of mathematics. Unfortunately, the mathematical representation and processing of the vagueness in language, which is one the main areas of study in FuSeTh, do not have a standard notation, and indeed, we found that the notation is very different in almost every document that one tries to read. This reason motivates us to make more of an effort to present an accurate notation to be used for the study of the concepts of the FuSeTh. The document is divided into four parts. The first one is devoted to the basic mathematical objects, concepts and notation related to the FuSeTh. In the other three parts, readers will find some mathematical ideas and some examples of Logic, measure theory and the theory of computation but from the FuSeTh perspective. The required prerequisites for this work are the basic concepts of the classical (crisp) set theory, crisp logic, measure theory and the theory of computation.

As a result of this work a paper, titled *Fuzzy sets. A way to represent ambiguity and subjectivity* arose and was published in Boletín de Matemáticas in 2017 [46]. We were motivated to give in this paper some of the basics notions of Fuzzy Mathematics, a branch of mathematics which has been in a continuous development for the latest fifty years. Also, we intend to present in the future more works on fuzzy mathematics focused on the understanding of fuzzy integrals, fuzzy derivatives and the topology of the set $\mathcal{F}(X)$, particularly $\mathcal{F}_C(\mathbb{R})$.

Chapter 1

Basic Notions

1.1 Fuzzy Sets

One of the most important mathematical notions which is frequently used in every day life is the concept of set [85]. Due to this, it is important to understand what a fuzzy set is. Fuzzy sets are based on **Crisp sets** (i.e., classical sets); the term "crisp" was introduced by Buckley to mean "not fuzzy" [9]. In this text, hereinafter, a crisp number means just a real number, a crisp matrix is a matrix whose elements are real numbers and a crisp solution to a problem is a solution involving crisp sets, crisp numbers, crisp functions and so on.

Example 1.1.1 (**Crisp Sets**)**.** *Let $X := [0, 10]$ be the referential universe. Then, the following are (crisp) subsets of X*

- $A := [3, 5]$.
- $B := \{2,\ 3,\ 5,\ 7\}$.
- $C := [0, 10] \cap \mathbb{N}$.
- $D := \left[0, \frac{\pi}{2}\right] \cup \left\{\frac{3}{2}\pi, \frac{5}{2}\pi\right\}$.

In order to define a fuzzy set, it is necessary to consider a referential universe and a function.

Definition 1.1.1. *Let X be a referential universe. We say that the pair $(A, \tilde{A}(x))$ is a* ***fuzzy (sub)set*** *of X if A is a subset of X endowed with a*

$$\begin{aligned} \tilde{A} : X &\longrightarrow [0, 1] \\ x &\longmapsto \tilde{A}(x) \end{aligned} \tag{1.1}$$

The function $\tilde{A}$ is usually called ***membership function*** *and also, in literature, it can be found denoted by μ_A. The* ***grade of membership*** *of x in the fuzzy set A is denoted by $\tilde{A}(x)$ and the* ***collection of all fuzzy subsets of*** *X by $\mathcal{F}(X)$.*

The collection $\mathcal{F}(X)$ is an analogous of the crisp power set, indeed, $\mathcal{F}(X) = [0, 1]^X$. The value $\tilde{A}(x)$ classify the element x as: A **total included member** (if $\tilde{A}(x) = 1$), a **non-included member** (if $\tilde{A}(x) = 0$) and a **fuzzy member** (if $0 < \tilde{A}(x) < 1$). To illustrate the previous definition and this observation, the following example is given.

Example 1.1.2 (Fuzzy Sets). *Let $X := [0, 10]$ be the referential universe. Then, the following are (fuzzy) subsets of X*

- *The pair $(A, \tilde{A}(x))$, where $A := [3, 5]$ and*

$$\tilde{A}(x) := \begin{cases} 0.3 & x = 3 \\ 0.2 & x \in (3, 4) \\ 0.7 & x = 4 \\ 0.5 & x \in (4, 5) \\ 0.901 & x = 5 \\ 0 & \textit{otherwise.} \end{cases}$$

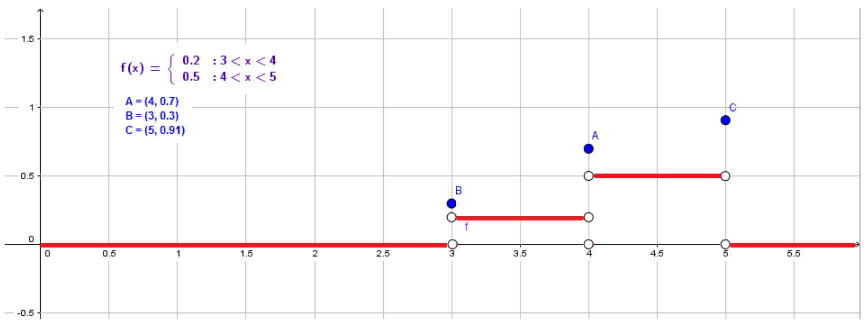

FIGURE 1.1. Graph of $\tilde{A}(x)$.

- *The pair $(B, \tilde{B}(x))$ where $B := \{2,\ 3,\ 5,\ 7\}$ and*

$$\tilde{B}(x) := \begin{cases} 1 & x \in B \\ 0 & \textit{in other case.} \end{cases}$$

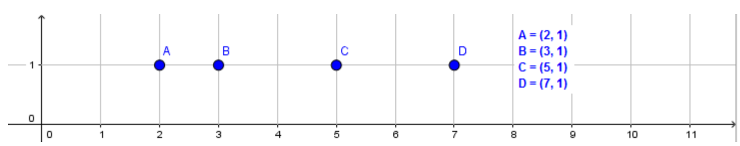

FIGURE 1.2. Graph of $\tilde{B}(x)$.

- *The pair $(C, \tilde{C}(x))$ where $C := [0, 10] \cap \mathbb{N}$ and*

$$\tilde{C}(x) := \begin{cases} 1 & x = 0 \\ \frac{1}{x} & (0, 10] \cap \mathbb{N} \\ 0 & \textit{in other case.} \end{cases}$$

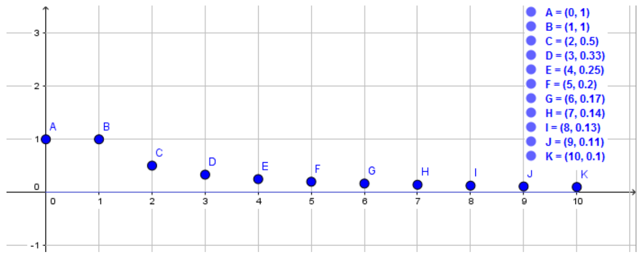

FIGURE 1.3. Graph of C(x).

- *The pair* $(D, \tilde{D}(x))$ *where* $D := \left[0, \frac{\pi}{2}\right] \cup \left\{\frac{3}{2}\pi, \frac{5}{2}\pi\right\}$ *and*

$$\tilde{D}(x) := \begin{cases} 1 & x \in [0, \frac{\pi}{2}) \\ |\sin(x)| & x \in D \setminus [0, \frac{\pi}{2}) \\ 0 & \textit{in other case.} \end{cases}$$

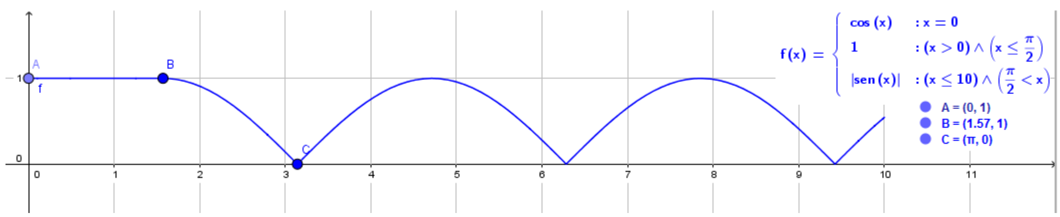

FIGURE 1.4. Graph of $\tilde{D}(x)$.

Remark* 1.** ***The domain of a fuzzy set can be any class and its codomain can be extended to any lattice or poset *$(L, \preceq)$, in this case the fuzzy set is termed $L-$**set**. When $L = [0, 1]$ is endowed with the linear order of the unit interval, the $L-$set is a "classical" fuzzy set. If $L = \{a + bi \mid a, b \in [0, 1]\}$ and it is endowed with the partial order*

$$a + bi \preceq c + di \Leftrightarrow a \leq c \text{ and } b \leq d,$$

the $L-$set is known as ***complex fuzzy set*** *[6, 68].*

Remark* 2.** *The notation $(A, A(x))$ was employed by Zadeh [79] in order to refer to a fuzzy subset A but $\{(x, \tilde{A}(x))\}_{x \in X}$ can be used too. Additionally, if it is wanted* ***to emphasize the cardinality of the referential universe *use the following notation:*

A finite set	*An enumerable set*	*A continuous set*
$\frac{\tilde{A}(x_1)}{x_1} + \cdots + \frac{\tilde{A}(x_n)}{x_n}.$	$\sum_{n \in \mathbb{N}} \frac{\tilde{A}(x_n)}{x_n}.$	$\int \frac{\tilde{A}(x)}{x}.$

In practice, the membership function might change from one person to another so, **the aim behind fuzzy sets is to capture the idea of partial membership**. What is really important is to represent correctly the knowledge provided by an expert and capture the meaning he intends to give to his own words but its accuracy is improved by trial and error [21, 31].

Remark 3. Fuzzy sets can be characterized by its membership function. *Hereinafter, fuzzy sets will be dealt as functions* $\tilde{A}: X \to [0,1]$*, where* X *is the referential universe. In this text, a fuzzy set will be denoted simply by* A *and, to refer to its membership function, the symbol* $\tilde{A}(x)$ *will be used. The empty set is defined by the map* $\tilde{\emptyset}(x) = 0$ *and we will denote it by* $\emptyset$ *and the total set* X *is defined by the map* $\tilde{X}(x) = 1$ *and will be denoted by* X*.*

Remark 4. Every crisp set is also a fuzzy set! *(see Example 1.1.2). So, it is natural to think that fuzzy sets are a kind of generalization of the crisp ones and the membership functions might be treated as generalizations of the traditional characteristic function of a crisp subset of* X*.*

The following are useful concepts to work with fuzzy sets.

Definition 1.1.2. *Let* A *be a fuzzy set of* X *and* $\alpha \in [0,1]$ *fixed, then:*
An $\alpha-$***cut*** *of* A *is defined as the crisp set*

$$A_\alpha := \{x \in X \mid \tilde{A}(x) \geq \alpha\}$$

and, a ***strong*** $\alpha-$***cut*** *of* A *as the crisp set*

$$A_{\alpha^+} := \{x \in X \mid \tilde{A}(x) > \alpha\}.$$

The ***support*** *of* A *is defined as the crisp set* A_{0^+} *and the* ***core*** *of* A *as the crisp set* A_1*.*

Note that we have that $A_\gamma \subseteq A_\alpha \subseteq A_\beta$ whenever $1 > \beta > \alpha > \gamma > 0$, as it can be seen in FIGURE 1.5 but this contention does not imply that the $\alpha-$cuts must be connected as it is shown in FIGURE 1.6

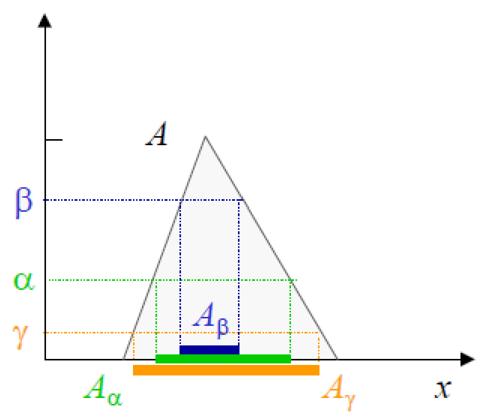

FIGURE 1.5. Fuzzy set $\alpha-$cuts.

The following example shows another difference between crisp sets and fuzzy ones.

Example 1.1.3. *Considering the sets given in Example 1.1.1 and Example 1.1.2, we have that in the crisp case the core, the support and all the* $\alpha-$*cuts coincide but, in the Fuzzy case:*

- ***The*** $0.5-$*cut* ***is:***
  $A_{0.5} = [4,5]$.

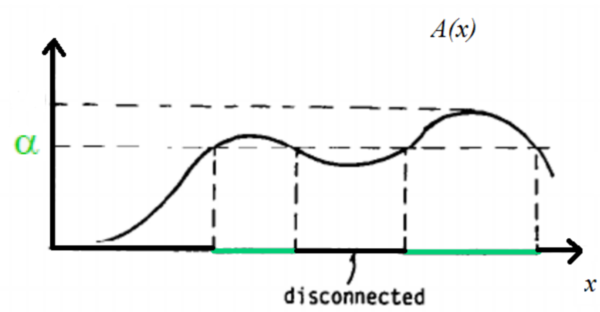

FIGURE 1.6. Non connected fuzzy set $\alpha-$cuts.

$B_{0.5} = \{2, 3, 5, 7\}$.
$C_{0.5} = \{0, 1, 2\}$.
$D_{0.5} = [0, \frac{5}{6}\pi] \cup [\frac{7}{6}\pi, \frac{11}{6}\pi] \cup [\frac{13}{6}\pi, \frac{17}{6}\pi] \cup [\frac{19}{6}\pi, 10]$.

- ***Support:***
  $A_{0^+} = [3, 5]$.
  $B_{0^+} = \{2,\ 3,\ 5,\ 7\}$.
  $C_{0^+} = [0, 10] \cap \mathbb{N}$.
  $D_{0^+} = [0, 10] \setminus \{\pi, 2\pi, 3\pi\}$.

- ***Core:***
  $A_1 = \emptyset$.
  $B_1 = \{2,\ 3,\ 5,\ 7\}$.
  $C_1 = \{0, 1\}$.
  $D_1 = [0, \frac{\pi}{2}] \cup \{\frac{3}{2}\pi, \frac{5}{2}\pi\}$.

1.2 Fuzzy Operators

From the membership functions, crisp set operations can be extended. Connectives over fuzzy sets are defined and studied through pointwise operations over the closed interval $[0, 1]$, but defining this kind of operators depends on the nature of the membership functions. The following definition presents the first fuzzy operations between fuzzy sets which were considered as generalizations of those in crisp sets. Afterwards, these operators were extended through the concepts of fuzzy negations, T-norms and T-conorms as we will present further.

Definition 1.2.1. *Let $A, B \in \mathcal{F}(X)$, the following are operations over fuzzy sets of X.*

- ***Fuzzy Inclusion*** *For all $x \in X$,*

$$\widetilde{A \subseteq B} := \tilde{A}(x) \leq \tilde{B}(x).$$

  The equality among two fuzzy subsets A, B is valid if and only if for all $x \in X$, $\tilde{A}(x) = \tilde{B}(x)$.

- ***Fuzzy Intersection*** *For all $x \in X$,*

$$\widetilde{A \cap B}(x) := \min\{\tilde{A}(x), \tilde{B}(x)\}.$$

- ***Fuzzy Union*** *For all* $x \in X$*,*

$$\widetilde{A \cup B}(x) := \max\{\tilde{A}(x), \tilde{B}(x)\}.$$

- ***Fuzzy Complementation*** *For all* $x \in X$*,*

$$\widetilde{A^c}(x) := 1 - \tilde{A}(x).$$

The following example can help the reader to understand the latest definition.

Example 1.2.1. *Considering again the sets given in Example 1.1.1 and Example 1.1.2, we have that in the* ***Crisp case****:* $B \subset C$*,* $A \cap B = \{3,5\}$ *and* $A \cup B = [3,5] \cup \{2,7\}$*. However, in the* ***Fuzzy case****:* $B \not\subseteq C$ *because* $\tilde{C}(2) = 0.5 < 1 = \tilde{B}(2)$*.*

- ***Intersection.***

$$\widetilde{A \cap B}(x) = \begin{cases} 0.3 & x = 3 \\ 0.901 & x = 5 \\ 0 & \textit{in other case.} \end{cases}$$

- ***Union.***

$$\widetilde{A \cup B}(x) = \begin{cases} 1 & x \in \{2,3,5,7\} \\ 0.2 & x \in (3,4) \\ 0.7 & x = 4 \\ 0.5 & x \in (4,5) \\ 0 & \textit{otherwise.} \end{cases}$$

We can observe that many properties for the crisp sets are preserved for fuzzy ones but the laws of contradiction and excluded middle (”tertio non datur”) do not hold for fuzzy sets [6], i.e.,

$$A \cap A^c \neq \emptyset \qquad \text{and} \qquad A \cup A^c \neq X.$$

For example, if we take in account the fuzzy set A from the Example 1.1.2, we can observe that $\emptyset \neq A \cap A^c$ because $\widetilde{A \cap A^c}(3) = 0.3$ and, $X \neq A \cup A^c$ since that $\widetilde{A \cup A^c}(3) = 0.7$.

In general, it is no difficult to see that are true for every fuzzy set:

$$0 \leq \widetilde{A \cap A^c}(x) = \min\{\tilde{A}(x), 1 - \tilde{A}(x)\} \leq 0.5$$

and

$$0.5 \leq \widetilde{A \cup A^c}(x) = \max\{\tilde{A}(x), 1 - \tilde{A}(x)\} \leq 1.$$

Additionally, there are operations that depend on the membership function. These operations are unary operations that modify the membership function of the Fuzzy Set, see FIGURE1.7. In practice, they are known as **hedges** and the most common are **Concentration type** and **Dilatation type**. The first one can be related to words like ’very ... ’ and reduces the membership grades like $\widetilde{A_C}(x) := \tilde{A}(x)^2$ does; the second one can be related to words like 'probably ...' and increases the membership grades like $\widetilde{A_D}(x) := \sqrt{\tilde{A}(x)}$ does.

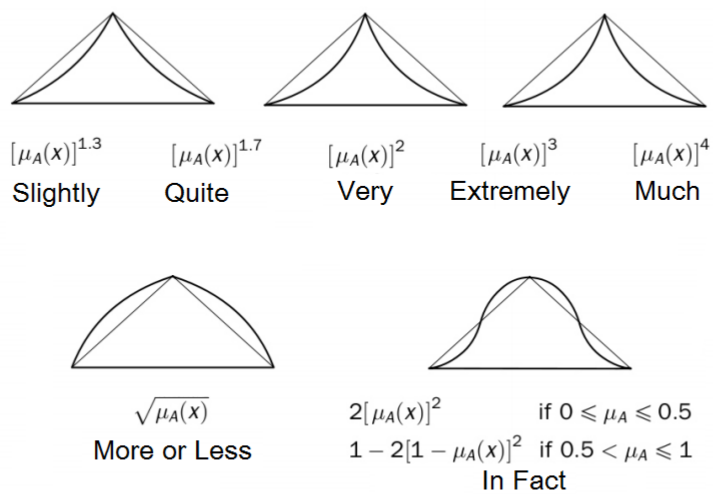

FIGURE 1.7. Taken from internet and modified.

By far, the most common operations are the those defined pointwise: the Fuzzy complement, the Fuzzy intersection and the Fuzzy union which are known as the **standard fuzzy operations**. The objective of this section is to present the main type of operators used on fuzzy sets which are: the N-complements, the T-norms, the T-conorms and the averaging operations. The first three types of operators extend the notions given previously for the complement, intersection (max) and union (min) operators, respectively. The last three ones are known as aggregation operators.

1.2.1 N-Complement

N-complements are also known as fuzzy negations, fuzzy complements and for a given fuzzy set, A, we denote this type of operator by A^c where c denotes a function, $N : [0,1] \to [0,1]$ which satisfy the following axioms [74]:

N(1) Membership dependency. The membership grade of x in the N-complement of A depends only on the membership grade of x in A. That is, for all $x \in X$:

$$\tilde{A}^c(x) = N(\tilde{A}(x)).$$

N(2) Boundary condition. The function of the N-complement, N, accomplishes $N(0) = 1$ and $N(1) = 0$.

N(3) Monotonicity The function N is monotonic nonincreasing, i.e., N fulfills that for all $a, b \in [0,1]$, if $a < b$ then $N(b) \leq N(a)$.

Also, it is desirable to consider the next two conditions:

N(4) Continuity. The function N is continuous.

N(5) Involution. The function N satisfies that for all $x \in [0,1]$, $N(N(x)) = x$.

When the function is strictly decreasing in condition **N(3)**, it is called a **strict negation** and if the condition **N(5)** is added then it becomes in a **strong negation**.

Some N-complements for the fuzzy set A are:

⋆ **Standard Negation**:

$$N(\tilde{A}(x)) = 1 - \tilde{A}(x).$$

N(1) Direct from the definition.

N(2) $N(0) = 1 - 0 = 1$ and $N(1) = 1 - 1 = 0$.

N(3) Let $0 \leq a < b \leq 1$, then $N(b) = 1 - b \leq 1 - a = N(a)$.

⋆ $\lambda-$**Complement (Sugeno complement)**: For $\lambda \in (-1, \infty)$,

$$N_\lambda(\tilde{A}(x)) = \frac{1 - \tilde{A}(x)}{1 + \lambda \tilde{A}(x)}.$$

N(1) Direct from the definition.

N(2) $N(0) = \frac{1-0}{1+\lambda 0} = \frac{1}{1} = 1$ and $N(1) = \frac{1-1}{1+\lambda} = \frac{0}{1+\lambda} = 0$.

N(3) Let $0 \leq a < b \leq 1$ and $\lambda \in (-1, \infty)$, then

$$\begin{array}{rcl} a(1+\lambda) & < & b(1+\lambda) \\ 1 - \lambda ab + a(1+\lambda) & < & 1 - \lambda ab + b(1+\lambda) \\ (1+\lambda a)(1-b) & < & (1+\lambda b)(1-a) \\ \dfrac{1-b}{1+\lambda b} & < & \dfrac{1-a}{1+\lambda a}. \end{array}$$

Thus, $N(b) \leq N(a)$.

⋆ **Yager complement**: For $w \in \mathbb{R}^+$,

$$N_w(x) = (1 - \tilde{A}(x))^{\frac{1}{w}}.$$

N(1) Direct from the definition.

N(2) $N(0) = (1-0)^{\frac{1}{w}} = 1^{\frac{1}{w}} = 1$ and $N(1) = (1-1)^{\frac{1}{w}} = 0^{\frac{1}{w}} = 0$.

N(3) Let $0 \leq a < b \leq 1$ and $w \in \mathbb{R}^+]$. By monotonicity of the function $n^{th}-$root we obtain $N(b) = (1-b)^{\frac{1}{w}} \leq (1-a)^{\frac{1}{w}} = N(a)$.

⋆ **Threshold complement**: For $t \in [0,1]$,

$$N_t(\tilde{A}(x)) = \begin{cases} 1 & \tilde{A}(x) \in [0,t] \\ 0 & \tilde{A}(x) \in [t,1]. \end{cases}$$

N(1) Direct from the definition.

N(2) Direct from the definition.

N(3) Let $0 \leq a < b \leq 1$ and $t \in [0,1]$. If $t < a < b$ then $N(b) = 0 \leq 0 = N(a)$. If $a < b < t$ then $N(b) = 1 \leq 1 = N(a)$ and $N(b) = 0 \leq 1 = N(a)$ when $a < t < b$.

An example is given below.

Example 1.2.2. *Let A be a fuzzy set whose membership function is defined as follows:*

$$\tilde{A}(x) = \begin{cases} \cos x & -\frac{\pi}{2} \leq x \leq 0 \\ 1 & 0 \leq x \leq \frac{\pi}{2} \\ \frac{|\frac{1}{2} - \sin x|}{2} & \frac{\pi}{2} \leq x \leq 5 \\ 0 & Otherwise. \end{cases}$$

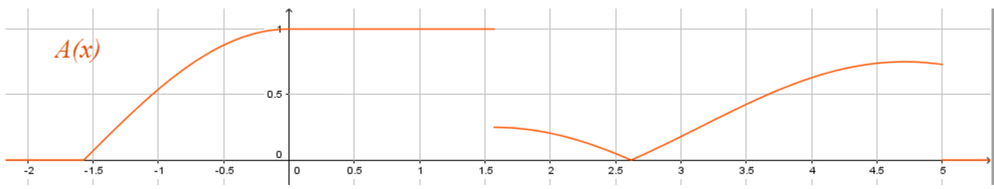

FIGURE 1.8. Fuzzy set A.

Then, its graph is shown in FIGURE 1.8 *and its standard negation, its Sugeno complement (for $\lambda = 1, -\frac{1}{2}, \frac{1}{\pi}$) and its Yager complement (for $\omega = 1, 2, \frac{1}{2}, \frac{1}{\pi}$) are shown in* FIGURE 1.9, FIGURE 1.10 *and* FIGURE 1.11, *respectively.*

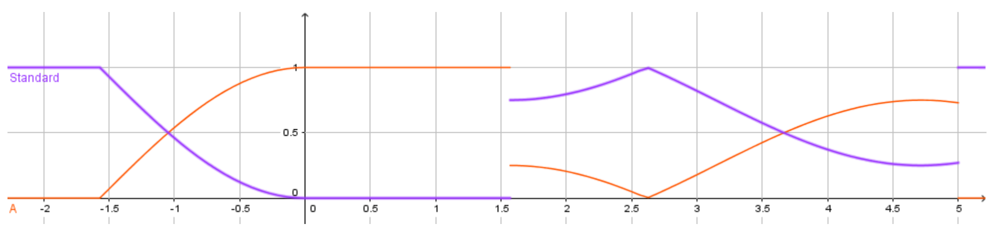

FIGURE 1.9. Standard negation.

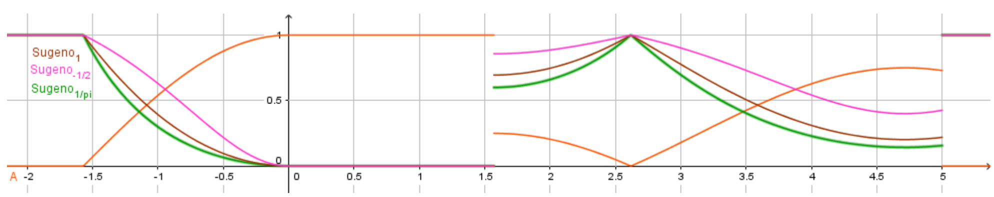

FIGURE 1.10. (Lambda) Sugeno complement.

1.2.2 Triangular norms (T-norm)

These type of operator are characterized by a function $T : [0,1] \times [0,1] \to [0,1]$ such that satisfies, for all $w, x, y, z \in [0,1]$:

T(1) Identity $T(1, w) = w$.

T(2) Monotonicity T satisfies that if $w \leq y$ and $x \leq z$ then $T(w, x) \leq T(y, z)$.

T(3) Commutativity $T(x, y) = T(y, x)$.

T(4) Associativity $T(x, T(y, z)) = T(T(x, y), z)$.

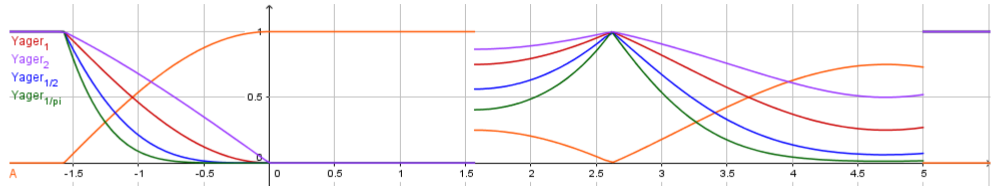

FIGURE 1.11. Yager Complement.

From the above conditions we can claim that

$$0 \leq T(x,0) = T(0,x) \leq T(0,1) = T(1,0) = 0 \qquad \text{and} \qquad T(1,1) = 1.$$

Given a fuzzy set A, some T-norms used for applications are:

⋆ **Standard**:

$$T(x,y) = \min\{x,y\}.$$

T(1) $T(1,\tilde{A}(x)) = \min\{1,\tilde{A}(x)\} = \tilde{A}(x)$.

T(2) Let $w \leq y$ and $x \leq z$ then $T(w,x) = \min\{w,x\} \leq \min\{y,z\} = T(y,z)$.

T(3) $T(w,x) = \min\{w,x\} = \min\{x,w\} = T(x,w)$.

T(4) $T(x,T(y,x)) = \min\{x,\min\{y,z\}\} \leq \min\{x,\min\{y,z\}\} = T(y,z)$. Just the following cases must be considered:

(1) $x \leq y \leq z$. (3) $y \leq x \leq z$. (5) $z \leq x \leq y$.
(2) $x \leq z \leq y$. (4) $y \leq z \leq x$. (6) $z \leq y \leq x$.

Here only the first case are shown because the others are similar.
$\min\{x,\min\{y,z\}\} = \min\{x,y\} = x = \min\{x,y\} = \min\{\min\{x,y\},z\}$.

⋆ **Drastic product**:

$$T_d(x,y) = \begin{cases} x & y = 1 \\ y & x = 1 \\ 0 & otherwise. \end{cases}$$

T(1) By definition of the function it is obtained that $T_d(1,\tilde{A}(x)) = \tilde{A}(x)$.

T(2) Let $A \in \mathcal{F}(X)$ and $w,x,y,z \in range(\tilde{A})$ such that $w \leq y$ and $x \leq z$. Then, just consider the following cases:

– $w \leq y = 1$ and $x = z = 1$:

$$T_d(w,x) = w \leq y = T_d(y,z).$$

– $w = y = 1$ and $x \leq z = 1$:

$$T_d(w,x) = x \leq z = T_d(y,z).$$

– $w \leq y < 1$ and $x \leq z \leq 1$:

$$T_d(w,x) = 0 \leq y = T_d(y,z).$$

– $w \leq y \leq 1$ and $x \leq z < 1$:

$$T_d(w, x) = 0 \leq z = T_d(y, z).$$

T(3) It is immediate due to $T_d(x, y)$ and $T_d(y, x)$ produce the same result, i.e.,

$$T_d(x,y) = \begin{cases} x & y = 1 \\ y & x = 1 \\ 0 & otherwise \end{cases} = T_d(y,x).$$

T(4) It is direct because $T_d(x, T_d(y, z))$ and $T_d(T_d(x, y), z)$ produce the same result, i.e.,

$$\begin{aligned} T_d(x, T_d(y,z)) &= \begin{cases} x & T_d(y,z) = 1 \\ T_d(y,z) & x = 1 \\ 0 & otherwise \end{cases} = \begin{cases} x & y = z = 1 \\ y & x = z = 1 \\ z & x = y = 1 \\ 0 & otherwise \end{cases} \\ &= \begin{cases} T_d(x,y) & z = 1 \\ z & T_d(x,y) = 1 \\ 0 & otherwise \end{cases} = \begin{cases} x & y = z = 1 \\ y & x = z = 1 \\ z & x = y = 1 \\ 0 & otherwise \end{cases} \\ &= T_d(T_d(x,y), z). \end{aligned}$$

⋆ **Algebraic product**:

$$T(x, y) = xy.$$

T(1) By definition of the function it is easily obtained that $T(1, \tilde{A}(x)) = \tilde{A}(x)$.

T(2) Let $0 \leq w \leq y$ and $0 \leq x \leq z$ then $T(w, x) = wx \leq yz = T(y, z)$.

T(3) $T(w, x) = wx = xw = T(x, w)$.

T(4) $T(x, T(y, x)) = T(x, yz) = xyz = T(xy, z) = T(T(x, y), z)$.

⋆ **Bounded difference**:

$$T_b(x, y) = \max\{0, x + y - 1\}.$$

T(1) $T_b(1, \tilde{A}(x)) = \max\{0, 1 + \tilde{A}(x) - 1\} = \max\{0, \tilde{A}(x)\} = \tilde{A}(x)$.

T(2) Let $0 \leq w \leq y$ and $0 \leq x \leq z$ then

$$T_b(x, w) = \max\{0, x + w - 1\} \leq \max\{0, y + z - 1\} = T_b(y, z).$$

T(3) $T_b(x, w) = \max\{0, x + w - 1\} = \max\{0, w + x - 1\} = T_b(w, x)$.

T(4) Let $x, y, z \in [0,1]$. Then $T_b(x, T_b(y,z))$ and $T_b(T_b(x,y),z)$ are equal since:

$$\begin{aligned} T_b(x, T_b(y,z)) &= \max\{0, x + T_b(y,z) - 1\} \\ &= \max\{0, x + \max\{0, y+z-1\} - 1\} \\ &= \begin{cases} \max\{0, x-1\} = 0 \\ \max\{0, x+y+z-2\} \end{cases} \\ &= \begin{cases} \max\{0, z-1\} = 0 \\ \max\{0, x+y+z-2\} \end{cases} \\ &= \max\{0, \max\{0, x+y-1\} + z - 1\} \\ &= \max\{0, T_b(x,y) + z - 1\} \\ &= T_b(T_b(x,y),z). \end{aligned}$$

Example 1.2.3. *Let B and C be a fuzzy set whose membership functions are defined by:*

$$B(x) = \begin{cases} \cos^8 x & -\frac{\pi}{2} \le x \le 0 \\ 1 - \sin^{\frac{1}{3}} x & 0 \le x \le \frac{\pi}{2} \\ 0 & Otherwise. \end{cases} \qquad C(x) = \begin{cases} -|\sin x|^{\frac{1}{2}} + 1 & -\frac{\pi}{2} \le x \le 0 \\ \cos^7 x & 0 \le x \le \frac{\pi}{2} \\ 0 & Otherwise. \end{cases}$$

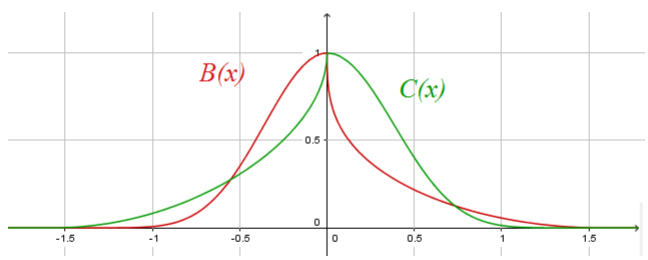

FIGURE 1.12. Fuzzy sets B and C.

Then, their graphs are shown in FIGURE 1.12 *and their standard, algebraic and bounded T-norms are shown in* FIGURE 1.13, FIGURE 1.14 *and* FIGURE 1.15 *respectively.*

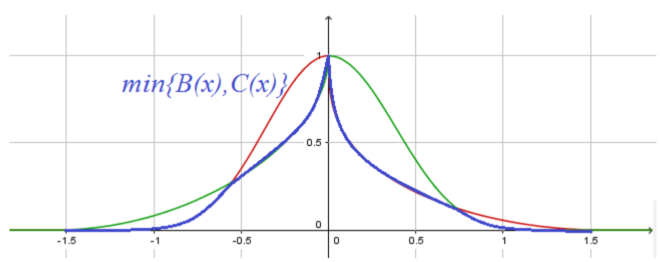

FIGURE 1.13. Standard T-norm, min.

In addition if T is any T-norm then for any $x, y \in [0,1]$:

$$T_{\min}(x,y) \le T(x,y) \le \min\{x,y\},$$

and the basic T-norms have the following order

$$T_{\min}(x,y) \le \max\{0, x+y-1\} \le xy \le \min\{x,y\}.$$

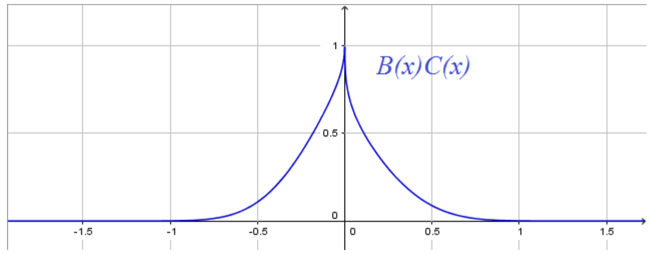

FIGURE 1.14. Algebraic T-norm.

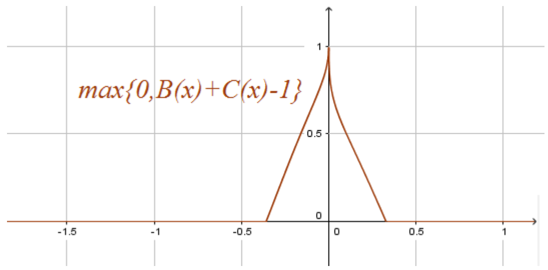

FIGURE 1.15. Bounded T-norm.

Since it is desirable to think in T-norms as binary operators, we denote T-norms by the symbol $\wedge$; so, $x \wedge y$ means $T(x, y)$.

1.2.3 Triangular conorms (T-conorms or S-norms)

These type of operator are characterized by a function $S : [0,1] \times [0,1] \to [0,1]$ such that for all $w, x, y, z \in [0,1]$:

S(1) Identity $S(w, 0) = w$.

S(2) Monotonicity S satisfies that if $w \leq y$ and $x \leq z$ then $S(w, x) \leq S(y, z)$.

S(3) Commutativity $S(x, y) = S(y, x)$.

S(4) Associativity $S(x, S(y, z)) = S(S(x, y), z)$.

From the above conditions we can claim that

$$1 = S(1,0) \leq S(1,x) = S(x,1) \leq 1 \qquad \text{and} \qquad S(0,0) = 0.$$

Some examples of T-conorms are:

⋆ **Standard**:

$$S(x, y) = \max\{x, y\}$$

S(1) $S(\tilde{A}(x), 0) = \max\{0, \tilde{A}(x)\} = \tilde{A}(x)$.

S(2) Let $w \leq y$ and $x \leq z$ then $S(w, x) = \max\{w, x\} \leq \max\{y, z\} = S(y, z)$.

S(3) $S(w, x) = \max\{w, x\} = \max\{x, w\} = S(x, w)$.

S(4) Let $x, y, z \in [0,1]$ then

$$\begin{aligned} S(x, S(y,x)) &= \max\{x, \max\{y,z\}\} \\ &= \begin{cases} \max\{x,y\} = \max\{y,z\} & x \le z \le y \text{ or } z \le x \le y \\ \max\{x,y\} = \max\{x,z\} & z \le y \le x \\ \max\{x,z\} = \max\{y,z\} & x \le y \le z \\ \max\{x,z\} & y \le x \le z \text{ or } y \le z \le x \end{cases} \\ &= \max\{\max\{x,y\}, z\} \\ &= S(S(x,y), z). \end{aligned}$$

⋆ **Drastic union**:

$$S_d(x,y) = \begin{cases} x & y = 0 \\ y & x = 0 \\ 0 & \textit{otherwise}. \end{cases}$$

S(1) By definition of the function it is obtained that $S_d(\tilde{A}(x), 0) = \tilde{A}(x)$.

S(2) Let $A \in \mathcal{F}(X)$ and $w, x, y, z \in \textit{range}(\tilde{A})$ such that $w \le y$ and $x \le z$. Then,

$$S_d(w,x) = \begin{cases} w & x = 0 \\ x & w = 0 \\ 0 & \textit{otherwise} \end{cases} \le \begin{cases} y & z = 0 \\ z & y = 0 \\ 0 & \textit{otherwise} \end{cases} = S_d(y,x).$$

S(3) It is immediate due to $S_d(x,y)$ and $S_d(y,x)$ produce the same result, i.e.,

$$S_d(x,y) = \begin{cases} x & y = 0 \\ y & x = 0 \\ 0 & \textit{otherwise} \end{cases} = S_d(y,x).$$

S(4) It is direct because $S_d(x, S_d(y,z))$ and $S_d(S_d(x,y), z)$ produce the same result, i.e.,

$$\begin{aligned} S_d(x, S_d(y,z)) &= \begin{cases} x & S_d(y,z) = 0 \\ S_d(y,z) & x = 0 \\ 0 & \textit{otherwise} \end{cases} = \begin{cases} x & y = z = 0 \\ y & x = z = 0 \\ z & x = y = 0 \\ 0 & \textit{otherwise} \end{cases} \\ &= \begin{cases} S_d(x,y) & z = 0 \\ z & S_d(x,y) = 0 \\ 0 & \textit{otherwise} \end{cases} = \begin{cases} x & y = z = 0 \\ y & x = z = 0 \\ z & x = y = 0 \\ 0 & \textit{otherwise} \end{cases} \\ &= S_d(S_d(x,y), z). \end{aligned}$$

⋆ **Algebraic sum**:

$$S(x,y) = x + y - xy.$$

S(1) By definition of the function it is easily obtained that $S(\tilde{A}(x), 0) = \tilde{A}(x) + 0 - \tilde{A}(x) \cdot 0 = \tilde{A}(x)$.

S(2) Let $0 \leq w \leq y$ and $0 \leq x \leq z$ then $S(w,x) = w+x-wx \leq y+z-yz = S(y,z)$.

S(3) $S(x,y) = x+y- = xy = y+x-yx = S(y,x)$.

S(4)

$$\begin{aligned} S(x,S(y,z)) &= x + S(y,z) - x \cdot S(y,z) \\ &= x+y+z-yz-x(y+z-yz) \\ &= x+y+z-yz-xy-xz+xyz \\ &= x+y-xy+z-z(x+y-xy) \\ &= S(S(x,y),z). \end{aligned}$$

⋆ **Bounded sum**:

$$S_B(x,y) = \min\{1, x+y\}.$$

S(1) $S_b(\tilde{A}(x), 0) = \min\{1, \tilde{A}(x)+0\} = \min\{1, \tilde{A}(x)\} = \tilde{A}(x)$.

S(2) Let $0 \leq w \leq y$ and $0 \leq x \leq z$ then

$$S_b(x,w) = \min\{1, w+x\} \leq \min\{1, y+z\} = S_b(y,z).$$

S(3) $S_b(x,y) = \min\{1, x+y\} = \min\{1, y+x\} = S_b(y,x)$.

S(4) Let $x,y,z \in [0,1]$. Then $S_b(x, S_b(y,z))$ and $S_b(S_b(x,y),z)$ are equal since:

$$\begin{aligned} S_b(x,S_b(y,z)) &= \min\{1, x+S_b(y,z)\} \\ &= \min\{1, x+\min\{1,y+z\}\} \\ &= \begin{cases} \min\{1, x+1\} \\ \min\{1, x+y+z\} \end{cases} \\ &= \begin{cases} \min\{1, 1+z\} \\ \min\{1, x+y+z\} \end{cases} \\ &= \min\{1, \min\{1, x+y\}+z\} \\ &= \min\{1, S_b(x,y)+z\} \\ &= S_b(S_b(x,y),z). \end{aligned}$$

Example 1.2.4. *Let B and C be a fuzzy set as in the Example 1.2.3. Then, the graphs of their standard, algebraic and bounded T-conorms are shown in* FIGURE 1.16, FIGURE 1.17 *and* FIGURE 1.18, *respectively.*

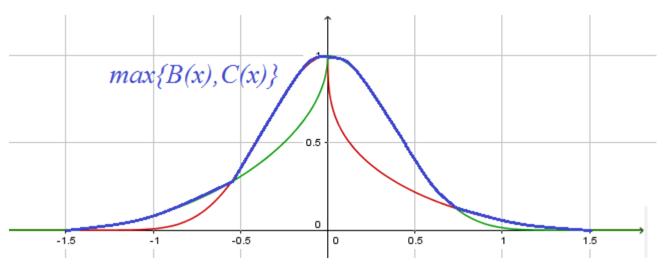

FIGURE 1.16. Standard T-conorm, max.

In addition if T is any T-norm then for any $x,y \in [0,1]$

$$\max(x,y) \leq S(x,y) \leq S_d(x,y),$$

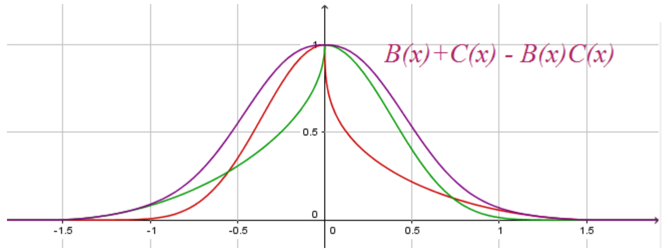

FIGURE 1.17. Algebraic sum.

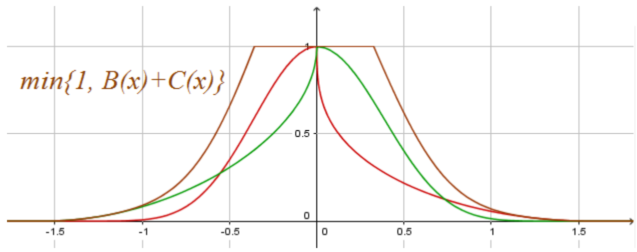

FIGURE 1.18. Bounded sum.

and the basic T-conorms have the following order

$$S_d(x, y) \geq \min\{1, x + y\} \geq x + y - xy \geq \max\{x, y\}.$$

Again, it is desirable to think in T-conorms as binary operators, then we denote T-conorms by the symbol $\vee$; so, $x \vee y$ means $S(x, y)$.

Triangular conorms are dual operator of T-norms and only certain combinations of T-norms, T-conorms and fuzzy complements satisfy the duality [74]. We say that a T-norm $\wedge$ and a T-conorm $\vee$ are dual with respect to the N-complement N if and only if

$$N(T(x, y)) = S(N(x), N(y))$$

and

$$N(S(x, y)) = T(N(x), N(y)).$$

For instance, the following T-norms and T-conorms are dual with respect to the standard fuzzy complement

- ⋆ $\min\{x, y\}$ and $\max\{x, y\}$.
- ⋆ ab and $a + b - ab$.
- ⋆ $\max\{0, x + y - 1\}$ and $\min\{1, x + y\}$.

1.2.4 Averaging Operator (Aggregation Operator).

These operators have the property of taking two arguments and produce a result greater than or equal to the $\min\{x, y\}$ and less than or equal to $\max\{x, y\}$.

An averaging operator is a function $h : [0, 1] \times [0, 1] \to [0, 1]$ such that for all $x, y, z, w \in [0, 1]$:

(A1) Idempotency $h(x, x) = x$.

(A2) Monotonicity $h(w, x) \leq h(y, z)$ whenever $w \leq y$ and $x \leq z$.

(A3) Commutativity $h(x, y) = h(y, x)$.

(A4) Continuity h is a continuous function.

Also, it is desired that such operators satisfy:

(A5) Extremes $h(0, 0) = 0$ and $h(1, 1) = 1$.

(A6) Boundary conditions $\min\{x, y\} \leq h(x, y) \leq \max\{x, y\}$.

The following are some examples of averaging operators:

Name	Operator
Arithmetic Mean	$\frac{x+y}{2}$
Generalized p-Mean	$\sqrt[p]{\frac{x^p+y^p}{2}}$
Harmonic Mean	$\frac{2xy}{x+y}$
Geometric Mean	$\sqrt{xy}$
Dual if Geometric Mean	$1 - \sqrt{(1-x)(1-y)}$

TABLE 1.1. Averaging operators.

Example 1.2.5. *Let B and C be a fuzzy set as in the Example 1.2.3. Then, the graphs of their arithmetic, generalized $p-$mean ($p = 2, 3, 5$), harmonic, geometric and dual IF-Geometric mean are shown in* FIGURE 1.19, FIGURE 1.20, FIGURE 1.21, FIGURE 1.22 *and* FIGURE 1.23, *respectively.*

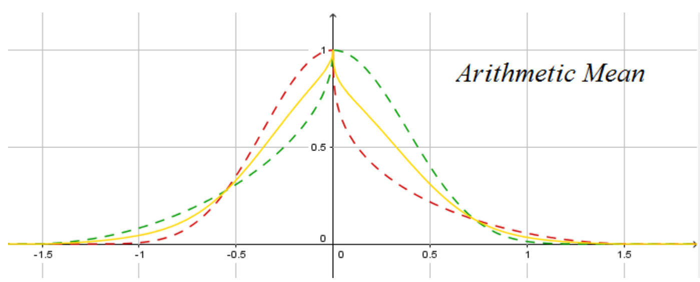

FIGURE 1.19. Arithmetic mean.

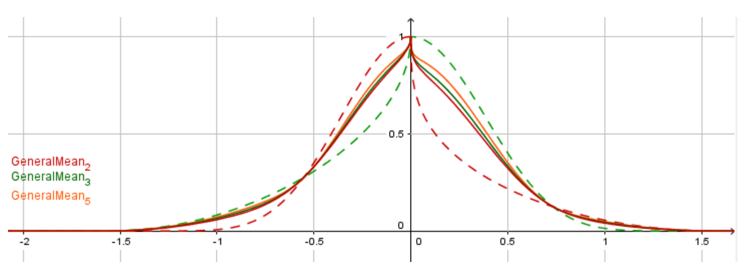

FIGURE 1.20. Generalized P-Mean.

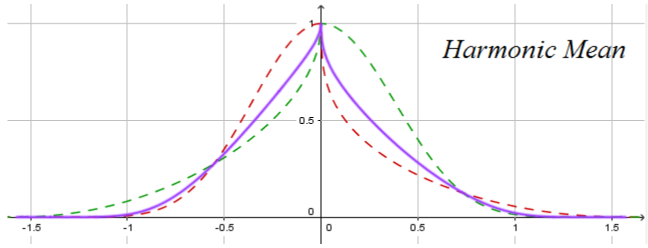

FIGURE 1.21. Harmonic Mean.

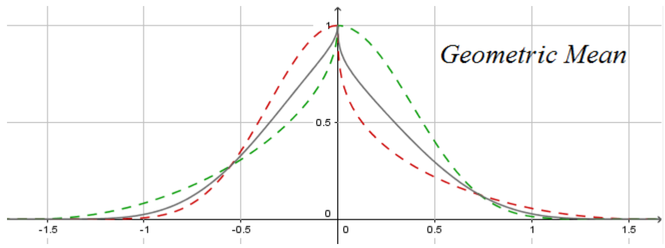

FIGURE 1.22. Geometric Mean.

There are other types of aggregation operators: the Ordered Weighted Averaging (OWA) operators and the Fuzzy measures and Fuzzy integrals. The first type of operators has been widely used in computational intelligence because of their ability to model linguistically expressed aggregation instructions. About the second type we can say that fuzzy measures and fuzzy integrals give a distinct treatment for aggregation and give rise to another class of aggregation operators which are used for computer vision, prediction, assessment of human reliability and, multi-attribute decision-making.

The Ordered Weighted Averaging (OWA) operator was introduced by Yager [75] and it is a weighted sum whose arguments are ordered. A formal definition is given as follows.

Definition 1.2.2. *Let $w = (w_1, \dots, w_n),\ w_i \in [0,1]$, be weights such that*

$$\sum_{i=1}^{n} w_i = 1$$

and let $\{\tilde{A}(x_i)\}$ be a finite sequence of increased ordered membership values of the fuzzy set A, i.e.,

$$\tilde{A}(x_1) \leq \cdots \leq \tilde{A}(x_n).$$

The OWA average of A is defined as

$$OWA_w(A) := \sum_{i=1}^{n} w_i \tilde{A}(x_i).$$

The OWA operator behaves as a compensatory operator similar to the generalized mean [48] and, although it seems very artificial, some special cases are:

1. $w = (1, 0, \dots, 0)$ then

$$OWA_w(A) = \tilde{A}(x_n) = \max_{1 \leq i \leq n} \tilde{A}(x_i).$$

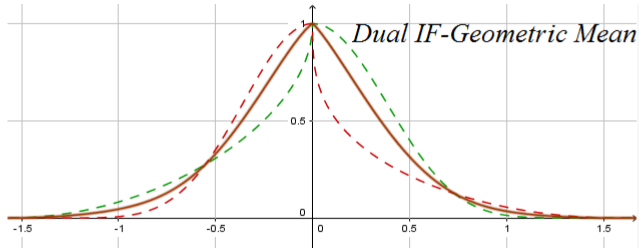

FIGURE 1.23. Dual IF-Geometric mean.

2. $w = (0, \ldots, 0, 1)$ then

$$OWA_w(A) = \tilde{A}(x_1) = \min_{1 \leq i \leq n} \tilde{A}(x_i).$$

3. $w = (\frac{1}{n}, \ldots, \frac{1}{n})$ then

$$OWA_w(A) = \frac{1}{n} \sum_{i=1}^{n} \tilde{A}(x_i).$$

The OWA operators are monotonic, bounded, continuous, symmetric and idempotent [48] and if the readers want to know more about the development of the OWA field, they are invited to consult the literature survey made by Emrouznejad [23].

Fuzzy measures come with an interesting interpretation in the context of sensor fusion and systematic diagnosis due to the sensor information is rarely precise and some uncertainty always remains [48]. We have devoted all the Chapter 2 to present the theory of fuzzy measures and present the notion of fuzzy integral which, in the context described above, is interpreted as a nonlinear aggregation of readings of the sensors.

Additionally, Zadeh proposed a way to extend crisp functions and it has become in an important tool for the FuSeTh and its applications. This extension principles are presented in the following subsection.

1.2.5 Zadeh's Extension Principle

Zadeh's extension principle serves for extending a real-valued functions into a corresponding fuzzy function [6].

Definition 1.2.3 (**First Zadeh's Extension Principle of** f)**.** *Given a function $f : X \longrightarrow Y$, where X and Y are crisp sets, it can be extended to a* ***(fuzzy) function*** *$F : \mathcal{F}(X) \longrightarrow \mathcal{F}(Y)$ such that $V = F(U)$, where*

$$V(y) := \begin{cases} \sup\{U(x) \mid x \in X \text{ and } f(x) = y\} & f^{-1}(y) \neq \emptyset, \\ 0 & \text{otherwise.} \end{cases} \tag{1.2}$$

Some examples are presented to clarify how the ZEP-1 works.

Example 1.2.6. *Let $X = \{a, b, c, d\}$ and $Y = \{1, 2, 3\}$. Now, consider the function $f : X \to Y$ given by $f(a) = f(b) = 1$, $f(c) = f(d) = 2$. Then the fuzzy set $B(x) = \frac{1}{a} + \frac{0.4}{b} + \frac{1}{c} + \frac{0.7}{d}$ in $\mathcal{F}(X)$ is extended to the fuzzy set $F(B) \in \mathcal{F}(Y)$ with values:*
- $F(B)(1) = \max\{B(x) \mid x \in f^{-1}(1)\} = \max\{B(a), B(b)\} = \max\{1,\ 0.4\} = 1,$

- $F(B)(2) = \max\{B(x) \mid x \in f^{-1}(2)\} = \max\{B(c), B(d)\} = \max\{1,\ 0.7\} = 1,$
- $F(B)(3) = 0$, *because* $f^{-1}(3) = \emptyset$.
Hence,

$$F(B) = \frac{1}{1} + \frac{1}{2},$$

that corresponds to the crisp subset $\{1,\ 2\}$ *of* Y.

Example 1.2.7. *Let* $f : [0, 10] \to [0, 1]$ *given by* $f(x) = 1 - \frac{x}{10}$ *and the fuzzy set* $A(x) = \frac{x}{10}$ *in* $\mathcal{F}([0, 10])$. *Then, for all* $z \in [0, 1]$,

$$F(A)(z) = \sup\{A(x) \mid f(x) = z\} = \sup\left\{\frac{x}{10} \mid 1 - \frac{x}{10} = z\right\} = 1 - z.$$

The second Zadeh's extension principle is a two dimensional case of the first one, i.e., it allows a crisp mapping $f : X \times Y \longrightarrow Z$, where X, Y, and Z are nonempty sets, to be extended to a mapping on fuzzy sets and it is very important because it allows us to extend operations between real numbers to the fuzzy case [6].

Definition 1.2.4 (**Second Zadeh's Extension Principle of** f)**.** *Given* $f : X \times Y \to Z$ *and* A *and* B *fuzzy sets, of* X *and* Y, *respectively, it is possible to build a function* $f : \mathcal{F}(X) \times \mathcal{F}(Y) \longrightarrow \mathcal{F}(Z)$ *by:*

$$g(A, B)(z) := \begin{cases} \sup\limits_{z=f(x,y)} \min\{A(x), B(y)\} & z \in Ranf, \\ 0 & z \notin Ranf. \end{cases} \tag{1.3}$$

***Remark* 5.** *Zadeh's extension is well defined for any fuzzy set* $A \in \mathcal{F}(X)$. *Indeed, when* $f^{-1}(y) \neq \emptyset$, *the set* $\{A(x) \mid x \in X, f(x) = y\}$ *is non-empty and bounded and so, it has a least upper bound [6].*

In the next section we give the most important example of a fuzzy subset: the fuzzy numbers. This type of fuzzy subsets of the set of the real numbers are defined, described and characterized below.

1.3 Fuzzy Numbers

Fuzzy numbers are special fuzzy subsets of the real numbers which are of great importance in fuzzy systems. That is why in this section the fuzzy numbers are defined, described and characterized. In applications, continuous fuzzy numbers are used and, among various shapes of them, triangular (shaped) fuzzy numbers and the trapezoidal (shaped) fuzzy numbers are the most popular ones.

Definition 1.3.1. *A* ***fuzzy number*** *is an element of* $\mathcal{F}(\mathbb{R})$ *whose membership function* $U : \mathbb{R} \to [0, 1]$ *satisfies the following:*

1. *Exists* $x_0 \in \mathbb{R}$ *such that* $U(x_0) = 1$. *(**Normality**).*

2. *Given* $x,\ y \in \mathbb{R}$ *and* $t \in [0, 1]$

$$U(tx + (1 - t)y) \geq \min\{U(x), U(y)\}. \ (\textbf{\textit{Fuzzy Convexity}}).$$

3. For any $x_0 \in \mathbb{R}$, it holds that

$$U(x_0) \geq \lim_{x \to x_0^{\pm}} U(x). \ (\textbf{\textit{Upper Semi-Continuity}}).$$

4. The crisp set

$$\overline{U_{0+}} = \overline{\{x \in \mathbb{R} | U(x) > 0\}}^{\mathbb{R}}$$

*is a compact set. (**Compact Support**).*

The set of all fuzzy numbers will be denoted by $\mathcal{F}_C(\mathbb{R})$

According to the definition of fuzzy number given above, the fuzzy set represented by the FIGURE 1.24 is a fuzzy number. To see an example of a fuzzy set which is not a fuzzy number see FIGURE 1.4 in **Example** 1.1.2.

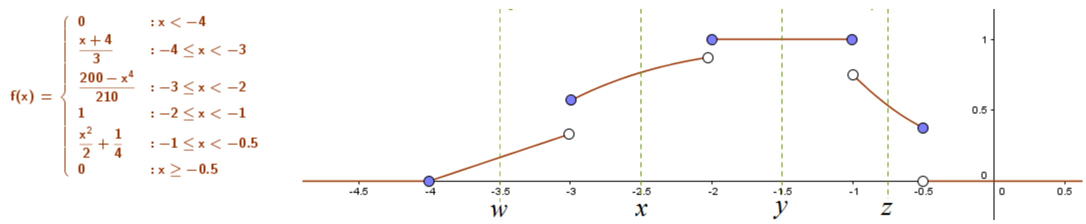

FIGURE 1.24. Fuzzy Number.

***Remark* 6.** *Other ways to define the upper-continuity are [19]:*

- *For any non-decreasing sequence $\alpha_1 \leq \alpha_2 \leq \cdots \leq \alpha_n \leq \ldots$ whose limit is α,*

$$U_\alpha = \lim_{n \to \infty} \bigcap_{i=1}^{n} U_{\alpha_i}. \text{ In particular,} U_\alpha = \bigcap_{\beta < \alpha} U_\beta.$$

- *For all $\epsilon > 0$, exists $\delta > 0$ such that if $|x - x_0| < \delta$ then $U(x) - U(x_0) < \epsilon$*

***Remark* 7.** ***Every real number is a fuzzy number**, i.e. $\mathbb{R} \subset \mathcal{F}_C(\mathbb{R})$, In effect, $\mathbb{R}$ can be identified as*

$$\mathbb{R} = \{\chi_x \mid x \in \mathbb{R}\} \qquad \text{where} \qquad \chi_x(y) := \begin{cases} 1 & x = y, \\ 0 & x \neq y. \end{cases}$$

Furthermore, fuzzy numbers generalize closed intervals and the following set may be taken into account:

$$\mathbb{I}_{\mathbb{R}} = \{\chi_{[a,b]} \mid [a, b] \text{ is an usual closed interval in } \mathbb{R}\}.$$

Before giving the characterizations of fuzzy numbers, the following lemma is presented.

Lemma 1.3.1. *If U is fuzzy convex, then U_α is convex for each $\alpha \in I$.*

Proof. [21] Let U be fuzzy convex and $x, y \in U_0$ for some $\alpha \in (0,1]$, so $U(x) \geq \alpha$ and $U(y) \geq \alpha$. Then, for any $\lambda \in [0,1]$,

$$U(\lambda x + (1-\lambda)y) \geq \min\{U(x), U(y)\} \geq \alpha.$$

So, $\lambda x + (1-\lambda)y \in U_\alpha$. Hence, for all $\alpha \in (0,1]$, U_α is a convex subset of $\mathbb{R}$. □

Now we are going to study some theorems used to characterize fuzzy numbers.

1.3.1 Characterization of Fuzzy Numbers Theorems

The following theorems are employed to characterize the fuzzy numbers and their proofs can be seen in the Bede's book [6].

Theorem 1.3.1 (Stacking Theorem). *If $U \in \mathcal{F}_C(\mathbb{R})$ is a fuzzy number and for $\alpha \in [0,1]$ the sets U_α are its $\alpha-$cuts, then:*

(i) U_α is a closed interval, i.e., for any $\alpha \in [0,1]$:

$$U_\alpha = [U_\alpha^-, U_\alpha^+],$$

where $U_\alpha^- := \inf U_\alpha$ and $U_\alpha^+ := \sup U_\alpha$.

(ii) If $0 \leq \alpha \leq \beta \leq 1$, then $U_\beta \subset U_\alpha$.

(iii) For any sequence α_n which converges from below to $\alpha \in (0,1]$ we have:

$$\bigcap_{n\geq 1} U_{\alpha_n} = U_\alpha.$$

(iv) For any sequence α_n which converges from above to 0 we have:

$$\overline{\bigcup_{n\geq 1} U_{\alpha_n}} = U_0.$$

Proof. Let $U \in \mathcal{F}_C(\mathbb{R})$ be a fuzzy number and U_α, for $\alpha \in [0,1]$, its $\alpha-$cuts, then:

(i) First, note that every set U_α is nonempty and bounded since $U_1 \neq \emptyset$ and the fact $\overline{U_{0^+}}$ is a compact set in $\mathbb{R}$ implies U_{0^+} is bounded. Let U be a fuzzy number and $\alpha \in (0,1]$. If $a, b \in U_\alpha$, then $U(a) \geq \alpha$ and $U(b) \geq \alpha$. Then from the fuzzy convexity, if $x \in [a,b]$ the $x \in U_\alpha$ since

$$U(x) \geq \min\{U(a), U(b)\} \geq \alpha.$$

As a conclusion U_α contains any closed interval $[a,b]$ and so U_α is a convex set. All is left to be proven is that U_α is closed.
From Upper Semicontinuity, if $U(x_0) < \alpha$ then there is an open interval W with $x_0 \in W$ such that $U(x) < \alpha$, for all $x \in W$. Then, the set $\{x \mid U(x) < \alpha\}$ is open and then its complement is a closed set, i.e., U_α is closed. Therefore, U_α is a closed interval for any $\alpha \in [0,1]$ because on the real line, closed convex sets are closed intervals.

(ii) if $0 < \alpha_1 \leq \alpha_2 \leq 1$ then, if $x \in U_{\alpha_2}$ then $U(x) \geq \alpha_2 \geq \alpha_1$ and so, $x \in U_{alpha_1}$. On the other hand, if $\alpha_1 = 0$ or $\alpha_2 = 0$, the the result is immediate.

(iii) Consider a non-decreasing sequence (α_n) such that converges to α. Then $U_{\alpha_n} \subseteq U_{\alpha_{n-1}}$, is a descending sequence of closed intervals $U_{\alpha_n} = [U^-_{\alpha_n}, U^+_{\alpha_n}]$. By the Nested Interval Theorem[63], we can achieve $U^-_{\alpha_n}, U^+_{\alpha_n}$ converge that is $U^-_{\alpha_n} \to a, U^+_{\alpha_n} \to b$ and consequently

$$[a, b] = \bigcap_{n\in\mathbb{N}} U_{\alpha_n}.$$

So, it is enough to show that $U(a) \geq \alpha$ and $U(b) \geq \alpha$. Suppose that $U(a) < \alpha$, then since U is upper semicontinuous, there is a neighbourhood W of a, such that $U(x) < \alpha$. This implies the existence of a rank $N \in \mathbb{N}$ with $U(U^-_{\alpha_n}) < \alpha$ for any $n \geq N$. Then since $\alpha_n \to r$ we obtain that there exists $n \in \mathbb{N}$ such that $U(U^-_{\alpha_n}) < \alpha_n$ which is a contradiction. Then it follows that $U(a) \geq \alpha$. Similarly we can show that $U(b) \geq \alpha$ so, $U(x) \geq \alpha$ and then $[a, b] \subseteq U_\alpha$. Additionally, from (ii), we have $U_\alpha \subseteq U_{\alpha_n}$ and it applies $U_\alpha \subseteq [a, b]$. Then finally we get $[a, b] = U_\alpha$, that is,

$$U_\alpha = \bigcap_{n\in\mathbb{N}} U_{\alpha_n}.$$

(iv) Since U_0 is a closed set and $\bigcup_{n\geq 1} U_{\alpha_n} \subseteq U_0$, we have that $\overline{\bigcup_{n\geq 1} U_{\alpha_n}} \subseteq U_0$. Reciprocally, $x \in U_0$ implies that there is a sequence $\{x_n\}_{n\in\mathbb{N}} \subseteq \{x \in \mathbb{R} \mid U(x) > 0\}$ that converges to x. Without loss of generality we may assume that $x_n \in U_{\alpha_n} \subseteq \bigcup_{n\geq 1} U_{\alpha_n}$. Then, we obtain $x \in \overline{\bigcup_{n\geq 1} U_{\alpha_n}}$.

So, we have completely proved the Stacking Theorem. □

The following theorem is the reciprocal of the Theorem 1.3.1.

Theorem 1.3.2 (**Negoita - Ralescu Characterization Theorem**). *Given a family of subsets of* $\mathbb{R}$, $\{M_\alpha\}_{\alpha\in[0,1]}$, *that satisfies the following conditions:*

(i) M_α *is a non-empty closed interval for any* $\alpha \in (0, 1]$.

(ii) If $0 \leq \alpha \leq \beta \leq 1$, *then* $M_\beta \subset M_\alpha$.

(iii) For any sequence α_n *which converges from below to* $\alpha \in (0, 1]$:

$$\bigcap_{n\geq 1} M_{\alpha_n} = M_\alpha.$$

(iv) For any sequence α_n *which converges from above to* 0:

$$\overline{\bigcup_{n\geq 1} M_{\alpha_n}} = M_{0^+}.$$

Then there exists a unique $U \in \mathcal{F}_C(\mathbb{R})$, *such that* $U_\alpha = M_\alpha$, *for any* $\alpha \in [0, 1]$.

Proof. [6] Let $\{M_\alpha\}_{\alpha\in[0,1]}$ a family of subsets of $\mathbb{R}$ that fulfills the properties $(i)-(iv)$. Define:

$$U(x) := \begin{cases} 0 & \text{if } x \notin M_{0^+} \\ \sup\limits_{x\in M_\alpha} \alpha & \text{if } x \in M_{0^+}. \end{cases}$$

Then U has the following features:

1. Since M_1 is nonempty, for some $x_0 \in M_1$ we have $U(x_0) = \sup\limits_{x_0\in M_\alpha} \alpha = 1$. Thus, U **is normal**.

2. Let $x \in [a,b] \subseteq M_{\underset{0}{+}}$ be a fixed element. Suppose that

$$U(a) = \sup_{a\in M_\alpha} \alpha = \alpha_a$$

and

$$U(b) = \sup_{b\in M_\alpha} \alpha = \alpha_b.$$

Note that $a \in M_{\alpha_a}$ and $b \in M_{\alpha_b}$. Now suppose that $\alpha_a \le \alpha_b$. Then, from the property (ii), we have $M_{\alpha_b} \subseteq M_{\alpha_a}$ and then $b \in M_{\alpha_a}$. Furthermore, for all $x \in (a,b)$, $\{\alpha \mid a \in M_\alpha\} \subseteq \{\beta \mid a \in M_\beta\}$. Hence,

$$U(x) \ge \alpha_a = U(a) = \min\{U(a), U(b)\}.$$

A symmetric reasoning can be followed in the case $\alpha_a \ge \alpha_b$. Thus, U **satisfies the fuzzy convexity.**

3. For all $\alpha \in (0,1]$, $U_\alpha = \{x \mid U(x) \ge \alpha\} = M_\alpha$.

In fact, let $\alpha_0 \in (0,1]$ be fixed and $x \in M_{\alpha_0}$ then $\alpha_0 \in \{\alpha \mid x \in M_\alpha\}$ and

$$U(x) = \sup_{x\in M_\alpha} \alpha \ge \alpha_0.$$

This implies that $x \in U_{\alpha_0}$. So, we have obtained $M_{\alpha_0} \subseteq U_{\alpha_0}$. For the symmetric inclusion we consider $x \in U_{\alpha_0}$, i.e., $U(x) \ge \alpha_0$. Now let us suppose that strict inequality $U(x) > \alpha_0$ holds. Then $\sup\limits_{x\in M_\alpha} \alpha > \alpha_0$ and there exists $\alpha_1 \ge \alpha_0$ with $x \in M_{\alpha_1}$. Since $M_{\alpha_1} \subseteq M_{\alpha_0}$, according to the property (ii), we obtain $x \in M_{\alpha_0}$ which completes the reasoning in this case.

If we suppose $U(x) = \sup\limits_{x\in M_\alpha} \alpha = \alpha_0$ then there exists a sequence (α_n) that converges from below to α_0 such that $x \in M_{\alpha_n}$, $n \ge 1$. From property (iii) we have

$$x \in \bigcap_{n\ge1} M_{\alpha_n} = M_{\alpha_0}.$$

As a conclusion we obtain $U_{\alpha_0} \subseteq M_{\alpha_0}$. So, $U_{\alpha_0} = M_{\alpha_0}$.

4. Since $U_\alpha = M_\alpha$ for all $\alpha \in [0,1]$, let (α_n) be a non-decreasing sequence such that $\alpha_n \to \alpha$. Observe that $M_{\alpha_{n+1}} \subseteq M_{\alpha_n}$ and then

$$U_\alpha = M_\alpha = \bigcap_{n\geq 1} M_{\alpha_n} = \bigcap_{n\geq 1} U_{\alpha_n} = \lim_{N\to\infty} \bigcap_{n=1}^{N} M_{\alpha_n}.$$

   Thus, U **is upper semicontinuous**.

5. It is easy to observe that $\overline{U_{0^+}} = M_{0^+}$. In fact,

$$\overline{U_{0^+}} = \overline{\{x \mid U(x) > 0\}} = \overline{\bigcup_{n\in\mathbb{Z}} \{x \mid U(x) \geq \alpha_n\}} = \overline{\bigcup_{n\geq 1} M_{\alpha_n}} = M_{0^+}.$$

   In addition, we obtain that U_{0^+} is a bounded subset of the real line. Then, it is compact.

Note that for all $x \in \mathbb{R}$, $0 \leq U(x) \leq 1$ and since U is a normal, fuzzy convex, upper semicontinuous and a compactly supported fuzzy set then we conclude that U is a fuzzy number whose $\alpha-$cuts are of the form $U_\alpha = M_\alpha$. □

The following is a characterization through monotonous functions.

Theorem 1.3.3 (**Representation Theorem.**). *Let U be a fuzzy number and let $U_\alpha = [U_\alpha^-, U_\alpha^+] = \{x \mid U(x) \geq \alpha\}$. Then the functions $L, R : [0,1] \longrightarrow \mathbb{R}$, defining the endpoints of the $\alpha-$cuts, satisfy the following conditions:*

(i) *$L(\alpha) := U_\alpha^- \in \mathbb{R}$ is a bounded, non-decreasing, left-continuous function in $(0,1]$ and it is right-continuous at 0.*

(ii) *$R := U_\alpha^+ \in \mathbb{R}$ is a bounded, non-increasing, left-continuous function in $(0,1]$ and it is right-continuous at 0.*

(iii) *$U_1^- \leq U_1^+$.*

Proof. For a given $U \in \mathcal{F}_C(\mathbb{R})$, and given $0 \leq \alpha \leq \beta \leq 1$, from the Stacking Theorem we obtain $U_\beta \subseteq U_\alpha$. Then we have, for all $0 \leq \alpha \leq \beta \leq 1$:

$$U_\alpha^- \leq U_\beta^- \leq U_1^- \leq U_1^+ \leq U_\beta^+ \leq U_\alpha^+.$$

Hence,

$$L(\alpha) \leq L(\beta) \text{ and } R(\beta) \leq R(\alpha).$$

therefore, the monotonicity properties and the third property are immediately satisfied.

Left continuity at $\alpha \in (0,1]$ follows from property (iii) of the Stacking Theorem. Indeed, let $\gamma \in (0,1]$ be fixed and (α_n) an increasing sequence converging to γ, i.e., $\alpha_n \to \gamma$. Then from the property (iii) of the Stacking Theorem it is obtained that

$$\bigcap_{n\geq 1} U_{\alpha_n} = U_\beta,$$

which immediately implies $U^-_{\alpha_n} \to U^-_\beta$ and $U^+_{\alpha_n} \to U^+_\beta$, i.e., $L(\alpha_n) \to L(\beta)$ and $R(\alpha_n) \to R(\beta)$. Thus, both functions are left continuous at arbitrary $\alpha_0 \in (0, 1]$.

In order to prove right continuity at 0, we consider a decreasing sequence (α_n) such that converges to 0, i.e., $\alpha_n \to 0$ and

$$\overline{U_{0^+}} = \overline{\{x \mid U(x) > 0\}} = \overline{\bigcup_{n\geq 1} \{x \mid U(x) \geq \alpha_n\}} = \overline{\bigcup_{n\geq 1} U_{\alpha_n}}.$$

□

The reciprocal of the Representation Theorem is presented in the next theorem.

Theorem 1.3.4 (**Goetschel - Voxman Characterization Theorem**). *Let* $L, R : [0, 1] \longrightarrow \mathbb{R}$ *be functions that satisfy the following conditions:*

(i) $L(\alpha) := U^-_\alpha \in \mathbb{R}$ *is a bounded, non-decreasing, left-continuous function in* $(0, 1]$ *and it is right-continuous at* 0*.*

(ii) $R(\alpha) := U^+_\alpha \in \mathbb{R}$ *is a bounded, non-increasing, left-continuous function in* $(0, 1]$ *and it is right-continuous at* 0*.*

(iii) $U^-_1 \leq U^+_1$.

Then there is a fuzzy number $U \in \mathcal{F}_C(\mathbb{R})$ *that has* U^-_α , U^+_α *as endpoints of its* $\alpha-$*cuts,* U_α*.*

Proof. [6] We will prove that the sets $M_\alpha = [U^-_\alpha, U^+_\alpha]$ satisfy the conditions of the Negoita-Ralescu characterization Theorem and so, this family of subsets defines a fuzzy number $U \in \mathcal{F}_C(\mathbb{R})$. From the monotonicity properties and from property (iii) we immediately obtain that for any $\gamma \leq \beta$,

$$U^-_\gamma \leq U^-_\beta \leq U^-_1 \leq U^+_1 \leq U^+_\beta \leq U^+_\gamma,$$

and this implies $M_\alpha = [U^-_\alpha, U^+_\alpha]$ are non-empty, closed intervals and that $M_\beta \subseteq M_\gamma$.

Let us consider now a sequence (α_n) which converges from below to $\alpha \in (0, 1]$. Then from the left continuity properties we obtain $U^-_{\alpha_n} \to U^-_\alpha$ and $U^+_{\alpha_n} \to U^+_\alpha$ and then we have

$$\bigcap_{n\geq 1} M_{\alpha_n} = M_\alpha.$$

From the right continuity at 0 we obtain

$$\overline{U_{0^+}} = \overline{\{x \in \mathbb{R} \mid U(x) > 0\}} = M_{0^+}.$$

Then there exists $U \in \mathcal{F}_C(\mathbb{R})$, such that $U_\alpha = M_\alpha = [U^-_\alpha, U^+_\alpha]$, for any $\alpha \in [0, 1]$. □

1.3.2 Types of Fuzzy Numbers

Several types of fuzzy numbers which are often used in applications are described below, we start by giving the most general fuzzy number then particular cases of them are presented.

Definition 1.3.2. *A* ***L-R Fuzzy Number*** *is a fuzzy set $U : \mathbb{R} \to [0,1]$ whose membership degree fulfills the following rule:*

$$U(x) = \begin{cases} 0 & x < a_0^- \\ L\left(\frac{x-a_0^-}{a_1^- - a_0^-}\right) & a_0^- \leq x < a_1^- \\ 1 & a_1^- \leq x < a_1^+ \\ R\left(\frac{a_{0^+}-x}{a_{0^+}-a_1^+}\right) & a_1^+ \leq x < a_{0^+} \\ 0 & a_{0^+} \leq x, \end{cases} \tag{1.4}$$

where $L, R : [0,1] \longrightarrow [0,1]$ are two continuous, increasing functions fulfilling $L(0) = R(0) = 0$, $L(1) = R(1) = 1$ and $a_0^- \leq a_1^- \leq a_1^+ \leq a_0^+$ are real numbers.

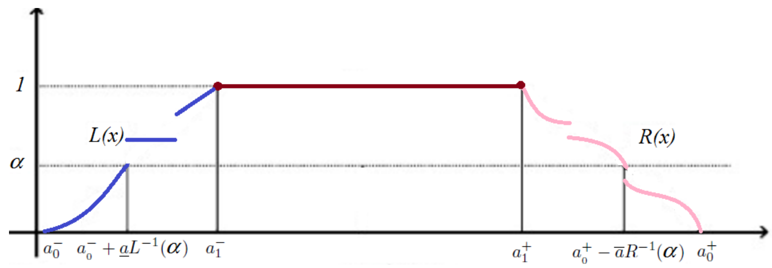

FIGURE 1.25. Graphical representation of a L-R fuzzy number.

The $\alpha-$cuts of a L-R Fuzzy Number are given by

$$U_\alpha = [a_0^- + \underline{a}L^{-1}(\alpha), a_0^+ - \overline{a}R^{-1}(\alpha)], \text{ where } \alpha \in [0,1]$$

where $\underline{a} := a_1^- - a_0^-$ and $\overline{a} := a_0^+ - a_1^+$. this values are called ***left spread*** *and* ***right spread****, respectively [6].*

L-R fuzzy numbers are considered important in the theory of fuzzy sets and they are very useful in applications. Symbolically, we write $U = (a_0^-, a_1^-, a_1^+, a_0^+)$, where $[a_1^-, a_1^+]$ is the core of U.

Particular cases of L-R fuzzy numbers are trapezoidal, triangular and Gaussian fuzzy numbers. Reader can see their shapes in FIGURE 1.26.

Definition 1.3.3. *A* ***Trapezoidal Fuzzy Number (TrFN)*** *is characterized by the membership function:*

$$U(x) = \begin{cases} 0 & t < a \\ \frac{t-a}{b-a} & a \leq t < b \\ 1 & b \leq t \leq c \\ \frac{d-t}{d-c} & c < t \leq d \\ 0 & d < t. \end{cases} \tag{1.5}$$

It is noted by four numbers $a < b < c < d$ where the base of the trapezoid is the interval $[a, d]$ and its top is over $[b, c]$ and the endpoints of the $\alpha-$cuts are given by

$$U_\alpha^- := a + \alpha(b-a) \quad \text{and} \quad U_\alpha^+ := d - \alpha(d-c).$$

Definition 1.3.4. *A* ***Triangular Fuzzy Number (TFN)*** *is characterized by the membership function:*

$$U(x) = \begin{cases} 0 & t < a \\ \frac{t-a}{b-a} & a \leq t < b \\ \frac{c-t}{c-b} & b < t \leq c \\ 0 & c < t. \end{cases} \tag{1.6}$$

It is noted by three numbers $a < b < c$ where the base of the triangle is the interval $[a, c]$ and its vertex is at $x = b$ and, the endpoints of the $\alpha-$cuts are given by

$$U_\alpha^- := a + \alpha(b - a) \quad \text{and} \quad U_\alpha^+ := c - \alpha(c - b).$$

Definition 1.3.5. *A* ***Gaussian Fuzzy Number (GFN)*** *is characterized by the membership function:*

$$U(x) = \begin{cases} 0 & x < x_1 - a\sigma_l \\ \exp\{-\frac{(x-x_1)^2}{2\sigma_l^2}\} & x_1 - a\sigma_l \leq x < x_1 \\ \exp\{-\frac{(x-x_1)^2}{2\sigma_r^2}\} & x_1 \leq x < x_1 + a\sigma_r \\ 0 & x_1 + a\sigma_r \leq x, \end{cases} \tag{1.7}$$

where x_1 is the core of the fuzzy number, σ_l, σ_r are the left and right spreads and $a > 0$ is a tolerance value.

Gaussian fuzzy numbers often are used in fuzzy control systems.

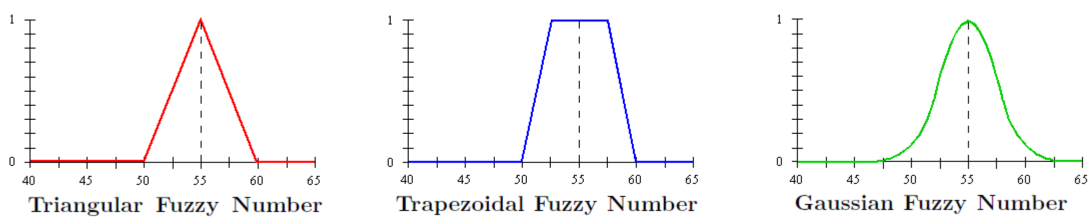

FIGURE 1.26. Types of fuzzy numbers used for applications.

Remark 8. *It does not matter if the universe of discourse is restricted to a closed compact interval, because all the fuzzy sets that fulfill normality, fuzzy convexity and upper-semicontinuity become in a fuzzy number.*

1.3.3 Defining an order in $\mathcal{F}_C(\mathbb{R})$

When the symbol $\leq$ appears, it should be clear which meaning is attached to it if the meaning is not stated but, in the context of the fuzzy sets, this symbol is used to indicate whether U is a fuzzy subset of V, i.e., $U(x) \leq V(x)$ for all $x \in \mathbb{R}$ [9]. In this section we will use some methods to define an order $U \leq V$, when U and V are fuzzy numbers.

Definition 1.3.6. *A* ***partial order*** $\leq$ *on* $\mathcal{F}_C(\mathbb{R})$ *must satisfy the following three properties:*

Reflexivity $U \leq U$.

Transitivity *$U \leq V$ and $V \leq W$, implies $U \leq W$.*

Skew-Symmetry *$U \leq V$ and $V \leq U$ implies $U \approx V$.*

If $\leq$ additionally satisfies that for U, V fuzzy numbers, $U \leq V$ or $U \leq V$ then, the relation $\leq$ is called a total (linear, complete) order on $\mathcal{F}_C(\mathbb{R})$.

The symbol $\leq$ is made up of two parts $<$ and $=$. Here, if U and V are fuzzy numbers, we will use $<$ to indicate that $U(x) \leq V(x)$ for all $x \in \mathbb{R}$ but we will not use $=$ because $U = V$ implies that the membership functions of U and V are identically equal. Instead of $=$ we will use $\approx$ to mean that the two fuzzy numbers are approximately equal or identical, i.e. $U(x) = V(x)$ for all $x \in X$ except a finite number of points. Thus, we have chosen the symbol $\lesssim$ to mean $U < V$ or $U \approx V$.

Let us now look at some definitions of $\lesssim$ on the set of fuzzy numbers.

Definition 1.3.7 (fuzzy-)maximum). *Let U and V be fuzzy numbers, by the Zadeh's Extension Principle:*

$$\widetilde{\max}\{U, V\} := \sup_{\max\{x,y\}=z} \min\{U(x), V(y)\}.$$

Note this is simply a fuzzification of $x \leq y$ if and only if $\max\{x, y\} = y$ for real numbers x and y.

Definition 1.3.8 (Hamming inequality). *Let U and V be fuzzy numbers, then $U \lesssim_H V$ is true whenever $d(V, \widetilde{\max}\{U, V\}) \leq d(U, \widetilde{\max}\{U, V\})$, where $d(U, V)$ is the **Haming distance of two fuzzy numbers** [6] U and V that belong to the space of continuous fuzzy numbers, i.e.,*

$$d(U, V) := \int_{-\infty}^{\infty} |U(x) - V(x)| dx.$$

Another definition for $U \lesssim V$ in terms of the $\alpha-$cuts is the next one.

Definition 1.3.9. *Let U and V be fuzzy numbers, $U_\alpha = [U_\alpha^-, U_\alpha^+]$ and $V_\alpha = [V_\alpha^-, V_\alpha^+)]$ their $\alpha-$cuts, then $U \lesssim_\alpha V$ is true if for every $\alpha \in [0, 1]$, $U_\alpha^- \leq V_\alpha^-$ and $U_\alpha^+ \leq V_\alpha^+$.*

1.4 Fuzzy Functions

Fuzzy functions are important in fuzzy modeling as crisp function are important in mathematical modeling [9]. Given referential universes X and Y, a fuzzy function F is simply a function mapping that maps elements from $\mathcal{F}(X)$ into elements of $\mathcal{F}(Y)$. The **set of all fuzzy functions from $\mathcal{F}(X)$ to $\mathcal{F}(Y)$** is denoted by $\Omega(X; Y)$.

1.4.1 How to get a fuzzy function?

An usual way to get a fuzzy function is to extend a crisp function to map fuzzy sets to fuzzy sets [33], and two methods to do it in the universe of fuzzy numbers are presented. The first method is based on the extension principle and the second one is based on the $\alpha-$cuts and the interval arithmetic.

The following result let us characterize how the fuzzy extension of a crisp real function f is through the $\alpha-$cuts of its argument.

Theorem 1.4.1 (ZEP-1 $\alpha-$cuts Theorem). *Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous function. It can be extended to a fuzzy function $F : \mathcal{F}_C(\mathbb{R}) \to \mathcal{F}_C(\mathbb{R})$ such that a given $U \in \mathcal{F}_C(\mathbb{R})$ can determine $V = F(U) \in \mathcal{F}_C(\mathbb{R})$ by means of its $\alpha-$cuts, i.e., $V_\alpha = f(U_\alpha)$, for all $\alpha \in [0,1]$.*

Observe that, for all $\alpha \in [0,1]$, U_α denotes the $\alpha-$cuts of U and:

$$V_\alpha = [V_\alpha^-, V_\alpha^+] = \left[\inf_{x\in U_\alpha} f(x), \sup_{x\in U_\alpha} f(x)\right].$$

Proof. First, let us prove that **if U_α and V_α are the $\alpha-$cuts of the fuzzy sets U and $V = F(U)$, respectively, then $V_\alpha = f(U_\alpha)$.**

The case $f^{-1}(y) = \emptyset$ is obvious. If $f^{-1}(y) \neq \emptyset$, by ZEP-1, we have $V = F(U)$ given by

$$V(y) := \sup_{\substack{x\in X\\ f(x)=y}} U(x).$$

Let $\alpha \in (0,1]$ fixed and $y \in f(U_\alpha)$, then there exist $x \in U_\alpha$ such that $y = f(x)$. Thus, $V(y) \geq U(x) \geq \alpha$ and $y \in V_\alpha$, i.e., $f(U_\alpha) \subseteq V_\alpha$. On the other hand, if $y \in V_\alpha$, $V(y) \geq \alpha$ and for any $\varepsilon > 0$ there exists an $x \in f^{-1}(y)$ such that $\alpha - \varepsilon \leq V(y) - \varepsilon \leq U(x)$. This implies that $U(x) \geq \alpha$. Hence, the Zadeh's extension of f satisfies that $V_\alpha \subseteq f(U_\alpha)$. Thus, $V_\alpha = f(U_\alpha)$ has been proved. In addition, if $V_\alpha = [V_\alpha^-, V_\alpha^+]$, we obtain that

$$V_\alpha^- = \inf_{x\in U_\alpha} f(x) \qquad \text{and} \qquad V_\alpha^+ = \sup_{x\in U_\alpha} f(x).$$

Now, let us prove now that **if U_α are the $\alpha-$cuts of a fuzzy number $U \in \mathcal{F}_C(\mathbb{R})$ then $V_\alpha = f(U_\alpha)$ define the $\alpha-$cuts of a fuzzy number $V \in \mathcal{F}_C(\mathbb{R})$ such that** $V = F(U)$. To do it, it will be proved that the family $\{V_\alpha\}_{\alpha\in(0,1]}$ satisfies the hypotheses of the Negoita-Ralescu characterization.

(i) Note that the U_α are compact convex intervals in $\mathbb{R}$ and, since f is continuous, we obtain that all the $V_\alpha = f(U_\alpha)$ are compact convex. Hence, V_α *is a closed interval for any* $\alpha \in [0,1]$.

(ii) If $\alpha \leq \beta$ then $U_\beta \subseteq U_\alpha$ and $V_\beta = f(U_\beta) \subseteq f(U_\alpha) = V_\alpha$.

(iii) Let us consider the sequence (α_n) that converges from below to $\alpha \in (0,1]$. Then, by Stacking Theorem,

$$\bigcap_{n\geq 1} U_{\alpha_n} = U_\alpha.$$

Hence,

$$V_\alpha = f(U_\alpha) = f\left(\bigcap_{n\geq 1} U_{\alpha_n}\right) \subseteq \bigcap_{n\geq 1} f(U_{\alpha_n}) = \bigcap_{n\geq 1} V_{\alpha_n}.$$

On the other hand, let $y \in \bigcap_{n\geq 1} V_{\alpha_n} = \bigcap_{n\geq 1} f(U_{\alpha_n})$. Then, for all $n \in \mathbb{Z}^+$, $y \in f(U_{\alpha_n})$. Thus, for all $n \in \mathbb{Z}^+$ there exists $x_n \in U_{\alpha_n}$ such that $y = f(x_n)$. But the sequence

(x_n) is a bounded one because $\{x_n\}_{n\in\mathbb{Z}^+} \subset \overline{U_{0^+}}$. Since U is a fuzzy number, $\overline{U_{0^+}}$ is compact crisp set, then there exists a subsequence (x_{n_k}) which converges. Let $x := \lim_{k\to\infty} x_{n_k}$ then $x \in U_\alpha$. Indeed, note that $U_\alpha \subseteq U_{\alpha_n}$ and if $x \notin U_\alpha$ then for some k, $x \notin U_{n_k}$ but, $x := \lim_{k\to\infty} x_{n_k}$, we obtain a contradiction. Thus, if $y \in \bigcap_{n\geq 1} V_{\alpha_n}$ then

$$y = \lim_{k\to\infty} y = \lim_{k\to\infty} f(x_{n_k}) = f\left(\lim_{k\to\infty} x_{n_k}\right) = f(x) \in f(U_\alpha) = V_\alpha.$$

Now, combining the two inclusions we obtain

$$\bigcap_{n\geq 1} V_{\alpha_n} = V_\alpha.$$

(iv) Let us consider a sequence (α_n) that converges from above to 0. Note that, $\overline{\bigcup_{n\geq 1} V_{\alpha_n}} \subseteq V_{0^+}$. On the other hand, since U is a fuzzy number, by the Stacking theorem, $\overline{\bigcup_{n\geq 1} U_{\alpha_n}} = U_{0^+}$. The continuity of f implies $f(\overline{A}) \subseteq \overline{f(A)}$ then

$$V_{0^+} = f(U_{0^+}) = f\left(\overline{\bigcup_{n\geq 1} U_{\alpha_n}}\right) \subseteq \overline{\bigcup_{n\geq} f(U_{\alpha_n})} = \overline{\bigcup_{n\geq 1} V_{\alpha_n}}.$$

Therefore, $V_0 = \overline{\bigcup_{n\geq 1} V_{r_n}}$

The hypotheses of Negoita-Ralescu characterization Theorem are fulfilled then V_α are $\alpha-$cuts of a unique fuzzy number. Since then the V_α are $\alpha-$cuts of V, it is concluded that $V \in \mathcal{F}_C(\mathbb{R})$. □

Also, it is possible to extend this idea to functions of many independent variables. Unfortunately, an extension principle based on compact convex subsets of $\mathbb{R}^2$ is not possible [6] but the following result allows us to extend real operations to the fuzzy number's case.

Theorem 1.4.2 (ZEP-2 $\alpha-$cuts Theorem). *Let $f : \mathbb{R}\times\mathbb{R} \to \mathbb{R}$ be a continuous function. Then it can be extended to a function $F : \mathcal{F}_C(\mathbb{R}) \times \mathcal{F}_C(\mathbb{R}) \to \mathcal{F}_C(\mathbb{R})$ such that, for any $U, V \in \mathcal{F}_C(\mathbb{R})$, the $\alpha-$cuts of $W := F(U, V)$ are of the form*

$$W_\alpha = \{f(x, y) \mid x \in U_\alpha, y \in V_\alpha\}$$

where, if $W_\alpha = [W_\alpha^-, W_\alpha^+]$ then

$$W_\alpha^- = \inf_{(x,y)\in U_\alpha\times V_\alpha} f(x, y) \qquad \text{and} \qquad W_\alpha^+ = \sup_{(x,y)\in U_\alpha\times V_\alpha} f(x, y).$$

Proof. First let us prove that if U_α, V_α and W_α are $\alpha-$cuts of the fuzzy sets U, V and $W := F(U, V)$, respectively, then $W_\alpha = f(U_\alpha, V_\alpha)$.

The case $f^{-1}(z) = \emptyset$ is easy to check. If $f^{-1}(z) \neq \emptyset$ then, by ZEP-2, $W = F(U, V)$ is given by

$$W(z) := \sup\{\min\{U(x), V(y)\} \mid x \in X, y \in Y, f(x, y) = z\}.$$

If $x \in U_\alpha$, $y \in V_\alpha$ then $U(x) \geq \alpha$, $V(y) \geq \alpha$ and, for $z := f(x,y)$, it implies also that $W(z) \geq \alpha$, i.e., $f(U_\alpha, V_\alpha) \subseteq W_\alpha$. On the other hand, if $W(z) \geq \alpha$ then for any $\varepsilon > 0$ there exists an $(x,y) \in f^{-1}(z)$ such that

$$W(z) - \varepsilon < U(x) \qquad \text{and} \qquad W(z) - \varepsilon < V(y)$$

which implies $U(x) \geq \alpha$ and $V(y) \geq \alpha$. So, we obtain $W_\alpha = f(U_\alpha, V_\alpha)$.

Since U_α, V_α are compact convex intervals in $\mathbb{R}$ and since f is continuous, we get $W_\alpha = f(U_\alpha, V_\alpha)$ compact convex.

If $\alpha \leq \beta$ then, $U_\beta \subseteq U_\alpha$ and $V_\beta \subseteq V_\alpha$. This implies $W_\beta = f(U_\beta, V_\beta) \subseteq f(U_\alpha, V_\alpha) = W_\alpha$.

Let now (α_n) be a sequence that converges from below to α and (β_n) converges from above to 0. Then, in a similar way as in the proof of Theorem 1.4.1, the relations

$$\bigcap_{n\geq 1} W_{\alpha_n} = W_\alpha \qquad \text{and} \qquad W_{0^+} = \overline{\bigcup_{n\geq 1} W_{\beta_n}}$$

can be obtained.

Since the hypotheses of Negoita-Ralescu characterization Theorem are fulfilled, finally we obtain $W \in \mathcal{F}_C(\mathbb{R})$. □

A fuzzy function derived from the extension principle will be denoted by F_E.

For all the functions we usually use in engineering and science we have an algorithm, using a finite number of additions, subtractions, multiplications and divisions, to compute the function with the desired accuracy [48]. Such functions can be extended to fuzzy functions by using $\alpha-$cuts and interval arithmetic.

Definition 1.4.1 ($\alpha-$**Cuts and Interval Arithmetic Procedure**). *Let $y = f(x)$ be such a function. We compute $\alpha-$cuts of $V = F(U)$ as*

$$V_\alpha = [V_\alpha^-, V_\alpha^+] := f([U_\alpha^-, U_\alpha^+]). \tag{1.8}$$

That is, we input the interval $U_\alpha := [U_\alpha^-, U_\alpha^+]$ into f and perform the arithmetic operations needed to evaluate f on this interval in order to obtain the interval $V_\alpha = [V_\alpha^-, V_\alpha^+]$. We must note that this does NOT necessarily mean

$$f(A_\alpha) = \{y \mid y = f(x), x \in A_\alpha\}.$$

For this procedure, only those crisp functions where the algorithm provides exact values with no approximation are considered. If the algorithm does not provide exact values for the function, then it is explicitly pointed out that we are using an approximation in the same way is done in the Buckley's book [9]. A fuzzy function derived from the Interval Arithmetic Procedure will be denoted by F_I.

***Remark* 9.** *For all the usual functions of engineering and science and for all continuous fuzzy numbers U in D interval of $\mathbb{R}$ we get*

$$F_E(U) \leq F_I(U).$$

A suggestion found given by Buckley [9] is always try the extension principle method first, and if it is too difficult to evaluate F_E then use F_I as an approximation to F_E.

1.4.2 Fuzzy Arithmetic

Let U and V be two fuzzy numbers then, it is natural to think about their arithmetic. Fuzzy numbers would be little use if there are no answer for questions related to these issues [74]. Fortunately, from the Zadeh's extension principle we can connect fuzzy sets with operations and tools of the classical mathematics.

Thus, to define the membership grade of x of the fuzzy arithmetic operations, from the ZEP, the following formula arises:

$$U(x) \star V(y) = \sup_{x \star y = z} \min\{U(x), V(y)\},$$

where $\star \in \{+, -, \times, \div, \vee(\max), \wedge(\min)\}$. It is necessary to keep in mind that these fuzzy operations are not just pointwise operations. The ZEP is a very general result which helps to define a fuzzy arithmetic of fuzzy numbers but it can be applied to many other situations, in fact, it can be applied to any crisp relation or function in mathematics to provide an analogous fuzzy one [74].

Since a fuzzy number U is characterized by having a closed interval as support, i.e. $U_0 = [U_0^-, U_0^+]$ and by using a quadruple $(U_0^-, U_1^-, U_1^+, U_0^+)$ to represent itself, it is possible to combine both ideas with interval arithmetic in order to give a best understanding for fuzzy arithmetic.

From the interval analysis we have the following definitions.

Definition 1.4.2 (**Arithmetic of Interval Operations**). *Given $[a, b]$ and $[c, d]$ in the set $\mathbb{I}_{\mathbb{R}}$ and $\lambda \in \mathbb{R}$,then [74]:*

1. $[a, b] + [c, d] := [a + c, b + d]$.
2. $[a, b] - [c, d] := [a - d, b - c]$.
3. $\lambda \cdot [a, b] = \begin{cases} [\lambda a, \lambda b] & \lambda \geq 0 \\ [\lambda b, \lambda a] & \lambda < 0 \end{cases}$
4. $[a, b] \times [c, d] := [\min A, \max A]$, *where* $A := \{ac, ad, bc, bd\}$.
5. $[a, b] \div [c, d] := [a, b] \times [\frac{1}{d}, \frac{1}{c}] = [\min B, \max B]$, *where* $B := \{\frac{a}{c}, \frac{a}{d}, \frac{b}{c}, \frac{b}{d}\}$.
   As long as zero does not belong to $[c, d]$ when we divide by this interval.

Hence, we can now characterize $W = U \star V$ through their $\alpha-$cuts, that is, for all $\alpha \in [0, 1]$, $W_\alpha = U_\alpha \star V_\alpha$, where $\star = \{+, -, \times, \div\}$ is an arithmetic operation between intervals. The following example explains how fuzzy arithmetic works.

Example 1.4.1. *Let $A, B \in \mathcal{F}_C(\mathbb{R})$ be TFN characterized by the following membership functions*

$$U(x) := \begin{cases} 0 & x \leq -1;\ x > 3 \\ \frac{x+1}{2} & x \in [-1, 1] \\ \frac{3-x}{2} & x \in (1, 3]. \end{cases} \qquad V(x) := \begin{cases} 0 & x \leq 1;\ x > 5 \\ \frac{x-1}{2} & x \in (1, 3] \\ \frac{5-x}{2} & x \in (3, 5]. \end{cases}$$

Graphically, this two fuzzy sets are

Their $\alpha-$cuts are of the form:

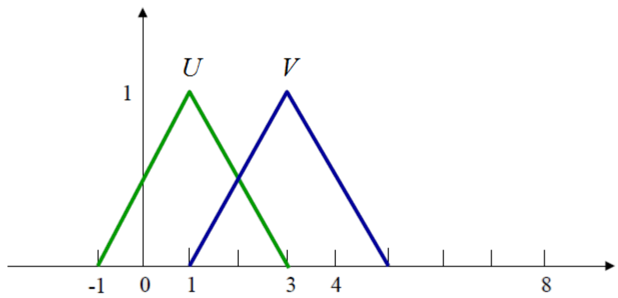

FIGURE 1.27. TFN A and B.

$$U_\alpha := [2\alpha - 1, 3 - 2\alpha] \qquad \textit{and} \qquad V_\alpha := [2\alpha + 1, 5 - 2\alpha].$$

Hence:

∗ *Addition is given by:*

$$U_\alpha + V_\alpha = [4\alpha, 8 - 4\alpha].$$

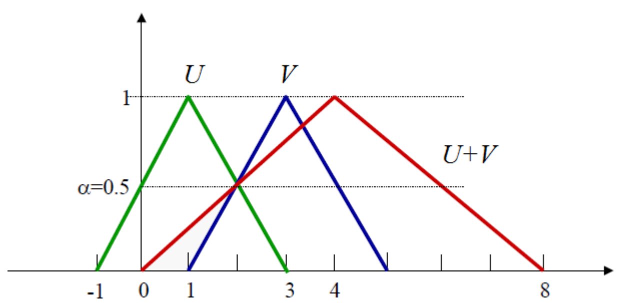

FIGURE 1.28. Adition of TFN.

∗ *Subtraction is given by:*

$$U_\alpha - V_\alpha = [4\alpha - 6, 2 - 4\alpha].$$

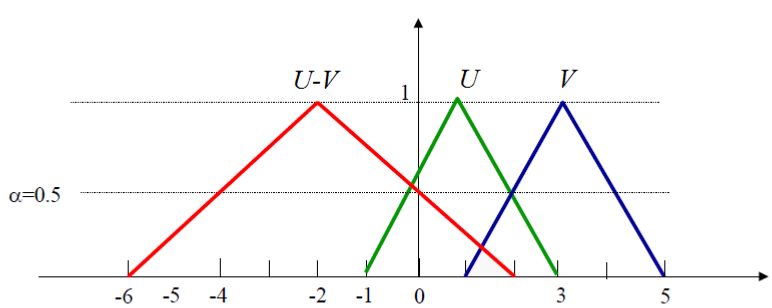

FIGURE 1.29. Subtraction of TFN.

∗ *Multiplication is given by:*

$$U_\alpha * V_\alpha = [-4\alpha^2 + 12\alpha - 5, 4\alpha^2 - 16\alpha + 15].$$

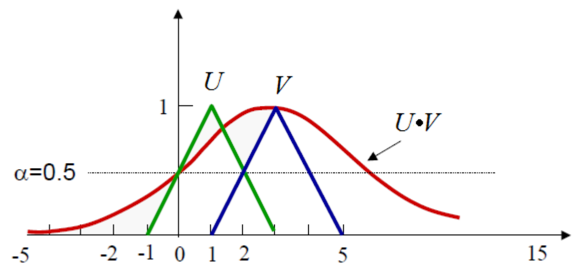

FIGURE 1.30. Multiplication of TFN.

∗ *and, division is:*

$$U_\alpha \div V_\alpha = \left[\frac{2\alpha - 1}{2\alpha + 1}, \frac{3 - 2\alpha}{2\alpha + 1}\right].$$

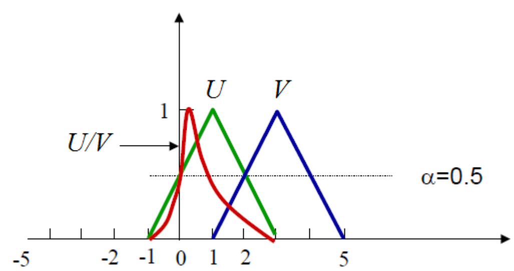

FIGURE 1.31. Division of TFN.

Remark* 10.** *Addition of fuzzy numbers satisfies commutativity, associativity, the existence of a neutral element but* ***None of $U \in \mathcal{F}_C(\mathbb{R}) - \mathbb{R}$ ***has an opposite in*** $\mathcal{F}_C(\mathbb{R})$ *(with respect to - w.r.t. - the operation +). Then, a conclusion is that* ***the space of fuzzy numbers is not a linear space****.*

The next definitions are alternative forms to define the subtractions and are very useful to define differentiability in $\mathcal{F}_C(\mathbb{R})$.

Definition 1.4.3. *The* ***Hukuhara difference (H-difference*** $\ominus_H$***)*** *is defined by*

$$U \ominus_H V = W \Longleftrightarrow U = V \dotplus W,$$

being $\dotplus$ *the standard fuzzy addition.*

The Hukuhara difference rarely exists, so several alternatives and generalizations were proposed like the generalized Hukuhara differentiability [6].

Definition 1.4.4 (($gH-$**difference**))**.** *Given two fuzzy numbers* $U, V \in \mathcal{F}_C(\mathbb{R})$*, the* ***generalized Hukuhara difference (*** $gH-$***difference)*** *is the fuzzy number* W*, if it exists, such that*

$$U \ominus_{gH} V = W \Longleftrightarrow U = V \dotplus W \text{ or } V = U - W.$$

Another type of difference used is the next one.

Definition 1.4.5 (($g-$**difference**)). *The **generalized difference** (**$g-$difference**) of two fuzzy numbers $U, V \in \mathcal{F}_C(\mathbb{R})$ is given by its $\alpha-$cuts as*

$$(U \ominus_g V)_\alpha = \overline{\bigcup_{\beta \geq \alpha} U_\beta \ominus_{gH} V_\beta},$$

for all $\alpha \in [0,1]$ where the $gH-$difference, $\ominus_{gH}$, is with interval operands U_β and V_β and it is well defined in this case.

We must note that for any fuzzy numbers $U, V \in \mathcal{F}_C(\mathbb{R})$ the $g-$difference $U \ominus_g V$ exists and it is a fuzzy number.

1.4.3 Types of Fuzzy Function

In this section all fuzzy sets and inputs to all fuzzy functions will be continuous fuzzy numbers. We will consider various types of elementary fuzzy functions. We again will use subscript "E" if the extension principle is used and subscript "I" when $\alpha-$cuts and interval arithmetic were used to fuzzify the crisp equation.

The following are some fuzzy functions which can be found in literature.

1. In order to fuzzify $y = f(x) = ax + b$, we can

   (a) make x fuzzy and have a and b crisp

   (b) make a and b fuzzy and have x crisp; or

   (c) make a, b and x fuzzy (the fully fuzzified linear function).

   The **fully fuzzified linear function** is

   $$Y = F(X) := A * X + B$$

   where X, Y, A, B are all continuous fuzzy numbers. If $X \neq A$, then $Y_E = Y_I$, but if $X = A$, then $Y = A^2 + B$ and Y_E may not equal Y_I [9].

2. The general **fully fuzzified polynomial of grade** n is

   $$Y = F(X) := \sum_{k=0}^{n} A_{n-k} * X^{n-k}$$

   for continuous fuzzy numbers $A_n, \cdots, A_0, X$. It is expected $Y_E \leq Y_I$ and $Y_E \neq Y_I$. **Fuzzy rational functions** are quotients of fuzzy polynomials.

3. The **fuzzy exponential** is also easy to compute because $y = \exp(x)$ is monotonically increasing. Consider for $A > 0$, $X > 0$ but $B < 0$

   $$Y = F(X) := A * \exp(B * X + C).$$

   Then $\alpha-$cuts of Y_E are $Y_{E\alpha} = [Y^-_{E\alpha}, Y^+_{E\alpha}]$, where

   $$Y^-_{E\alpha} = A^-_\alpha \exp(B^-_\alpha X^+_\alpha + C^-_\alpha)$$

and

$$Y_{E\alpha}^{+} = A_{\alpha}^{+} \exp(B_{\alpha}^{+} X_{\alpha}^{-} + C_{\alpha}^{+}).$$

4. The **fuzzy logarithmic function** is

$$Y = F(X) = A * \ln(B * X + C),$$

is also easy to evaluate since $y = \ln(x)$, $x > 0$, is monotonically increasing. To simplify things assume that $B > 0$, $X > 0$, $C > 0$ so that $B * X + C \geq 1$ and then $\ln(B * X + C) \geq 0$. Then $\alpha-$cuts of Y are

$$Y_{\alpha}^{-} = A_{\alpha}^{-} \ln(B_{\alpha}^{-} X_{\alpha}^{-} + C_{\alpha}^{-})$$

and

$$Y_{\alpha}^{+} = A_{\alpha}^{+} \ln(B_{\alpha}^{+} X_{\alpha}^{+} + C_{\alpha}^{+}).$$

In this case we have $Y_{E} = Y_{I}$.

5. Let X be a continuous fuzzy number. Let F and G be fuzzy functions then F **and** G **are inverses of each other**, written $G = F^{-1}$, if for all X for which these equations are defined

$$F(G(X)) = X \qquad \text{and} \qquad G(F(X)) = X.$$

Other special fuzzy functions may be discussed like fuzzifying $f(x) = |x|$, $\sqrt[n]{x}$, etc. In particular, since $\sqrt[2]{x}$, $x \geq 0$ is monotone increasing we easily find that $\sqrt[2]{X_{\alpha}} = [\sqrt[2]{X_{\alpha}^{-}}, \sqrt[2]{X_{\alpha}^{+}}]$ where $X_{\alpha} = [X_{\alpha}^{-}, X_{\alpha}^{+}]$.

Also it is natural to want to fuzzify functions like $y = \sin(x)$, $y = \cos(x)$ and $y = \tan(x)$ where it is used radians (real numbers) for x since that is what is needed in calculus. So, let X be a continuous fuzzy number and set $Y = F(X) := \sin(X)$ (the other trigonometric functions work similarly [9]).

From the extension principle we may find the $\alpha-$cuts of $Y = \sin(X)$ as follows:

$$Y_{\alpha} = [Y_{\alpha}^{-}, Y_{\alpha}^{+}] = [\min_{x \in X_{\alpha}} \sin(x), \max_{x \in X_{\alpha}} \sin(x)]$$

for $0 \leq a \leq 1$. Since $y = \sin(x)$ is periodic with period 2π, it is possible toobtain very different results for Y depending on the size of the support and core of X [9].

We can not expect crisp trigonometric identities like $\sin^2(x)+\cos^2(x) = 1$ and $\tan^2(x)+1 = \sec^2(x)$ to hold for fuzzy trigonometric functions. For instance, $\sin^2(X) + \cos^2(X) \neq 1$ because the left side of this equation is a fuzzy subset of the real numbers and the right side is not fuzzy, it is the crisp number one. But it is expected that number one belongs to the core of $\sin^2(X) + \cos^2(X)$ [9]. Although the fuzzy trigonometric identities do not hold exactly, the fuzzy trigonometric functions are still periodic and we have [9]:

$$\begin{aligned}
\sin(X + 2\pi) &= \sin(X), \\
\cos(X + 2\pi) &= \cos(X), \\
\tan(X + \pi) &= \tan(X).
\end{aligned}$$

1.5 Fuzzy Relations

Relations are important in mathematics as well as in the real world since they provide a way to capture and describe interactions, associations or interconnectedness between variables or elements of two or more sets. They are involved in logic, approximate reasoning, classification, rule-based systems, pattern recognition and control. On the other hand, fuzzy relations generalize the concept of relation by admitting the notion of partial association and are instrumental in problems of information retrieval, pattern classification, control and decision-making [48, 62].

A relation R is defined over the Cartesian product of X and Y can be represented by a mapping $R : X \times Y \to \{0, 1\}$ such that if x and y are related, $R(x, y) = 1$ and $R(x, y) = 0$ otherwise. In the fuzzy case, a **fuzzy relation on** $X \times Y$ is represented through a function $\tilde{R} : X \times Y \to [0, 1]$ such that for some pair (x, y), $\tilde{R}(x, y) = 0$ means that these elements are unrelated, if $\tilde{R}(x, y) = 1$ then they are fully related and when $0 < \tilde{R}(x, y) < 1$ it means that x and y are partially associated. We use $\tilde{R}$ to indicate that a relation R is a fuzzy one; also we say that $\tilde{R} = \{(x, y), \tilde{R}(x, y)\}_{(x,y)\in X\times Y}$.

The **domain** of a fuzzy relation $\tilde{R}$ defined on $X \times Y$ is a fuzzy subset whose membership function is $dom[R](x) = \sup\limits_{y\in Y} R(x, y)$ while its **codomain** has as membership function to $cod[R](x) = \sup\limits_{x\in X} R(x, y)$. A fuzzy relation can be represented through their $\alpha-$cuts too like $\tilde{R} = \bigcup\limits_{\alpha\in[0,1]} \alpha R_\alpha$ or, in terms of the membership function $\tilde{R}(x, y)$ as $\tilde{R}(x, y) = \sup\limits_{\alpha\in[0,1]} \min\{\alpha, R(x, y)\}$. Additionally, given $\tilde{A} \in \mathcal{F}(X)$ and $\tilde{B} \in \mathcal{F}(Y)$ then $\tilde{A}\times\tilde{B}$ is a fuzzy relation $\tilde{R}$ contained in $X \times Y$ whose membership function $\tilde{R}(x, y) = \min\{\tilde{A}(x), \tilde{B}(y)\}$.

Two fuzzy relations $\tilde{R}$ and $\tilde{S}$ defined over the same Cartesian product $X \times Y$ are **equal** if and only if for all $(x, y) \in X \times Y$, $\tilde{R}(x, y) = \tilde{S}(x, y)$ and we say that R **is included in** S if and only if for all $(x, y) \in X \times Y$, $\tilde{R}(x, y) \le \tilde{S}(x, y)$; it is denoted $R \subseteq S$. The **shadows** or projections of a fuzzy relation $\tilde{R}$ on X and Y are, respectively, the fuzzy sets defined upon X and Y with the membership functions given by

$$\widetilde{R_X}(x) = \sup_{y\in Y} \tilde{R}(x, y) \text{ and } \widetilde{R_Y}(y) = \sup_{x\in X} \tilde{R}(x, y).$$

Also, we can define operations among two fuzzy relations through their membership functions as follows:

Union

$$\widetilde{R \cup S}(x, y) = \max\{\tilde{R}(x, y), \tilde{S}(x, y)\}.$$

Intersection

$$\widetilde{R \cap S}(x, y) = \min\{\tilde{R}(x, y), \tilde{S}(x, y)\}.$$

Complement

$$\widetilde{R^c}(x, y) = 1 - \tilde{R}(x, y).$$

Containment

$$R \,\tilde{\subseteq}\, S \Longrightarrow \tilde{R}(x, y) \le \tilde{S}(x, y).$$

Transpose

$$\widetilde{R^T}(y,x) = \tilde{R}(x,y).$$

As a consequence of the definitions provided above: $(\widetilde{R^T})^T = \tilde{R}$ and $(\widetilde{R^c})^T = (\widetilde{R^T})^c$.

Let $\tilde{R}$, $\tilde{S}$ and $\tilde{T}$ be fuzzy relations on $X \times Y$, $Y \times Z$ and $X \times Z$, respectively, we define two types of composition.

Fuzzy max **-** min **composition** ($\tilde{T} = \tilde{S} \circ \tilde{R}$)

$$\tilde{T}(x,z) = \max_{y \in Y} \min\{\tilde{R}(x,y), \tilde{S}(y,z)\}.$$

Fuzzy max **- product composition** ($\tilde{T} = \tilde{S} \odot \tilde{R}$)

$$\tilde{T}(x,z) = \max_{y \in Y} \tilde{R}(x,y)\tilde{S}(y,z).$$

***Remark* 11.** *The composition $\tilde{R}$ and $\tilde{S}$ is similar to the classic product of the matrices, the difference is that we use* max *and* min *instead of the addition and the multiplication [62].*

A fuzzy relation on $X \times X$, $\tilde{R}$ is:

Reflexive Whether for all $x \in X$,

$$\tilde{R}(x,x) = 1.$$

Symmetrical If for all $x, y \in X$,

$$\tilde{R}(x,y) = \tilde{R}(y,x).$$

$\sup - \min$ **Transitive** When for all $x, y, z \in X$,

$$\sup_{z \in Z} \min\{\tilde{R}(x,z), \tilde{R}(z,y)\} \leq \tilde{R}(x,y).$$

Irreflexive Whether there exists $x \in X$ such that $\tilde{R}(x,x) \neq 1$.

Anti-reflexive Whether for all $x \in X$, $\tilde{R}(x,x) \neq 1$.

$\varepsilon-$**reflexive** Whether for all $x \in X$, $\tilde{R}(x,x) \geq \varepsilon$.

Asymmetric If there exist $x, y \in X$ such that $\tilde{R}(x,y) \neq \tilde{R}(y,x)$.

Antisymmetric When $\tilde{R}(x,y) > 0$ and $\tilde{R}(y,x) > 0$ implies that $x = y$.

Similarity (or a Fuzzy Equivalence Relation (FER)) whether $\tilde{R}$ is a reflexive, symmetric and $\sup - \min$ transitive fuzzy relation.

Compatibility (or proximity relation) if $\tilde{R}$ is reflexive and symmetric.

Fuzzy Pre-Order If $\tilde{R}$ is reflexive and $\sup - \min$ transitive.

Fuzzy Order If $\tilde{R}$ is reflexive, antisymmetric and $\sup - \min$ transitive.

***Remark* 12.** *We can extend the previous definitions by replacing the operator* min *by another* $t-$*norm in the* max − min *composition and the* sup − min *transitivity [68].*

Fuzzy relations can be represented through both fuzzy matrices and fuzzy graphs. A **fuzzy graph** $\tilde{G}$ is a pair $\tilde{G} = (V, E)$ where V is the set of **vertices** (nodes or elements) and E is the set of **fuzzy edges** which are elements of the fuzzy set $E : X \times Y \to [0, 1]$. In other words, a fuzzy graph is essentially a weighted graph whose weights belong to the unit interval. A fuzzy graph is frequently expressed through a fuzzy matrix.

Example 1.5.1. *Let* $\tilde{R}$ *be a fuzzy relation on* $X \times Y$ *defined by the matrix*

$$\begin{pmatrix} 0 & 0.3 & 0.4 \\ 0.2 & 0.5 & 0.3 \\ 0.8 & 0 & 0 \\ 0.7 & 0.7 & 0.9 \end{pmatrix}$$

has as graph

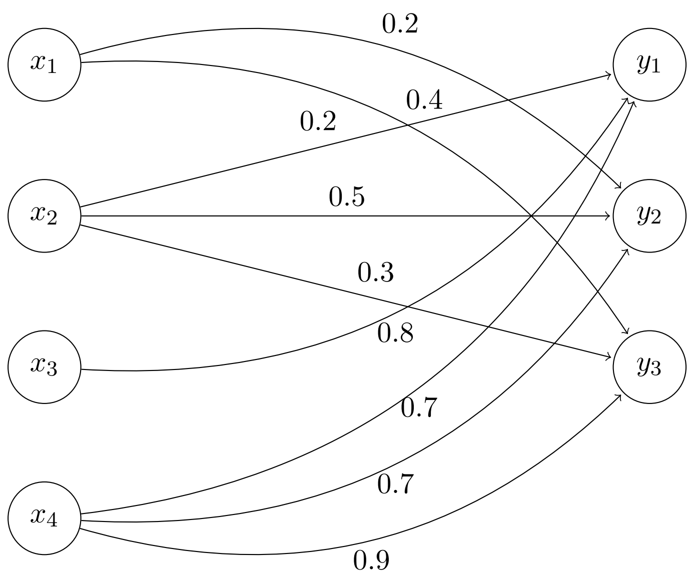

FIGURE 1.32. Graph of the fuzzy relation $\tilde{R}$.

1.6 Fuzzification and Defuzzification

Fuzzification is a kind of scientific permissiveness [88], it is the process of changing a real scale value into a fuzzy value (that is to give a membership function to a crisp set) and it is usually done by experience and analysis; so, a wrong fuzzification of the input variables might cause instability and error of the modeled system [69]. Types of fuzzifiers are **singleton fuzzifier, characteristic fuzzifier, Gaussian fuzzifier, trapezoidal fuzzifier, triangular fuzzifier**.

A fuzzyfication function f is a function that transforms every real number within the range of a given variable into a fuzzy quantity (fuzzy number) that approximates the real number. The following are some fuzzification approaches:

- **Gaussian fuzzification**: Let $a, r \in \mathbb{R}$, $a > 0$, $A(x) = \exp\{-\frac{x-r}{a}\}^2$.
- **Triangular fuzzification**: Let $b, r \in \mathbb{R}$, $b > 0$, $A = (r - b, r, r + b)$.

The motivation of fuzzification is that it may be more realistic to replace a crisp measurement x by a function which means "*around* x" in considering the measurement uncertainty due to the degrading sensors, environment disturbances or other reasons [71].

Since the results in fuzzy inferences are fuzzy sets which present an imprecise description of the system, defuzzification is thought to reduce such a fuzzy sets to "summarizes" the fuzzy set. To evaluate a defuzzification operator the following criteria are suggested [71]:

Let A, B be fuzzy sets of a subset X of $\mathbb{R}$.

1. **Core Selection**: $D(A) \in A_1$.
2. **Scale Invariance**: $D(aA + B) = D(A)$ where $(aA + B)(x) = aA(x) + b$ (Interval Scale Invariance). $D(aA) = D(A)$ (Ratio Scale Invariance). $D(A + b) = D(A)$ (Relative Scale Invariance).
3. **Monotonicity** $B(D(A)) = A(D(A))$, $B(x) \leq A(x)$ for all $x < D(A)$ and $B(x) \geq A(x)$ for all $x > D(A)$ imply $D(B) \geq D(A)$.
4. $T-$**conorm Criteria**: $D(A) \leq D(B)$ implies $D(A) \leq D(S(A, B))D(B)$ where S is a $T-$conorm.

The most frequently used one is the center of gravity method (also known as centroid method) however it just satisfy the monotonicity property described above [71].

On the other hand, **defuzzify** consists of replacing the fuzzy variable for a crisp one [50]. We can use different techniques, the most useful are:

- **Maximum Defuzzification Technique (MDT):** This method gives the output with the highest membership function. If the fuzzy set has membership function $A(x)$ then the picked element x^* satisfies for all x in the universe:

$$\mu(x^*) \geq \mu(x).$$

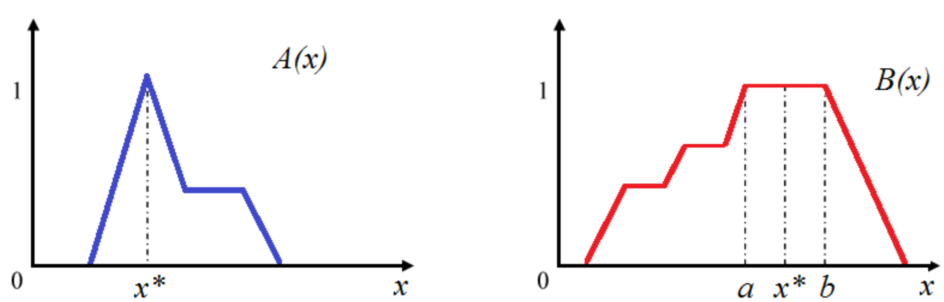

FIGURE 1.33. a is the smallest of max, b is the largest of max and x^* is the mean of max.

- **Centroid Defuzzification Technique (CoG - Center of Gravity):** This method purposed by Sugeno in 1985 is the most commonly used technique because it is very accurate [50]. If $A \in \mathcal{F}(X)$ then

$$x^* = \frac{\displaystyle\int_{U_{0+}} A(x)x dx}{\displaystyle\int_{U_{0+}} A(x) dx}.$$

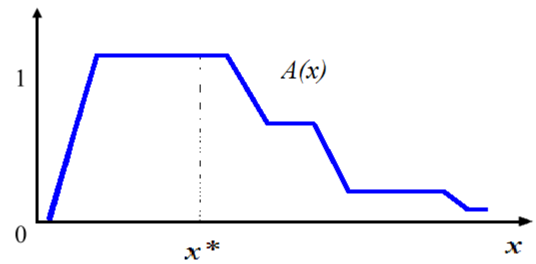

FIGURE 1.34. Centroid defuzzication point.

- **Weighted Average Defuzzification Technique (WAD):** It is other of the most frequently used in fuzzy applications since it is one of the more computationally efficient methods. Unfortunately, it is usually restricted to symmetrical output membership functions [50]. We can defuzzify by doing:

$$x^* = \frac{\sum A(x)x dx}{\sum A(x) dx}$$

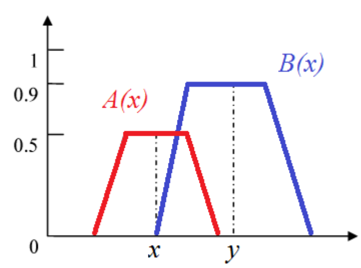

FIGURE 1.35. Weighted average defuzzication point for the fuzzy sets A and B.

- **Center of Area Defuzzification Technique (CoA):**. If $A \in \mathcal{F}(X)$ then this number is defined as the point of the support of A that divides the area under the membership function into two equal parts. If that number is a then it satisfies:

$$\int_{-\infty}^{a} A(x) dx = \int_{a}^{\infty} A(x) dx.$$

- **Expected value and expected interval (EV and EVI)**. If U is a continuous fuzzy number then the expected value is given by

$$EV(U) := \frac{1}{2}\int_0^1 (U_\alpha^- + U_\alpha^+) dr.$$

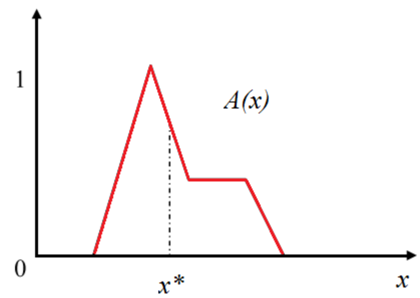

FIGURE 1.36. Center of Area defuzzication point.

and the expected interval is

$$EVI(U) := \left[\int_0^1 U_\alpha^- \, dr, \int_0^1 U_\alpha^+ \, dr \right].$$

The expected value is the midpoint of the expected interval [6].

1.7 Some Notation

In this section we present some abbreviations we use in the document and the more common symbols used to represent some of the concepts in the Fuzzy Set Theory.

FuSeTh .... Fuzzy Set Theory

ZEP .... Zadeh's Extension Principle

TFN .... Triangular Fuzzy Number

TrFN .... Trapezoidal Fuzzy Number

GFN .... Gaussian Fuzzy Number

MDT .... Maximum Defuzzification Technique

CoG .... Centroid Defuzzification Technique

WAD .... Weighted Average Defuzzification Technique

CoA .... Center of Area Defuzzification Technique

$A,\ B,\ C,\ \ldots$ .... Crisp Set

$\tilde{A},\ (A, \tilde{A}(x)),\ \{x, \tilde{A}(x)\}$ .... Fuzzy Set

$\sum_{x \in X} \frac{\tilde{A}(x)}{x}$ .... A is a Countable Fuzzy Set

$\int_{x \in X} \frac{\tilde{A}(x)}{x}$ .... A is an Uncountable Fuzzy Set

$\tilde{A}(x),\ \mu_A(x),\ A(x)$ .... Membership function of the Fuzzy Set A

$A_\alpha,\ [A]^\alpha$ .... $\alpha-$cut of the Fuzzy Set A

A_{0^+} .... Support of the Fuzzy Set A

A_1 .... Core of the Fuzzy Set A

$\mathcal{F}(X)$ .... Collection of all Fuzzy Subsets of X

$A \in \mathcal{F}(X)$ .... A is a Fuzzy Subset of X

$N(\cdot)$ .... $N-$Complement, Fuzzy Complement, Fuzzy Negation

$T(\cdot,\cdot)$ .... $T-$norm, norm

$S(\cdot,\cdot)$ .... $S-$norm, conorm

$U = (U_0^-,\ U_1^-,\ U_1^+,\ U_0^+)$ .... $L-R$ Fuzzy Number

$\mathcal{F}_C(\mathbb{R}),\ \mathbb{R}_F$ .... Collection of the all Fuzzy Numbers

$U \in \mathcal{F}_C(\mathbb{R})$ .... U is a Fuzzy Number

$U_\alpha = [U_\alpha^-,\ U_\alpha^+]$ .... $\alpha-$cut of a Fuzzy Number

U_α^- .... $\inf U_\alpha$

U_α^+ .... $\sup U_\alpha$

$\tilde{F},\ \tilde{G},\ \tilde{H},\ \ldots$ .... Fuzzy Function

$\tilde{R},\ \tilde{S},\ \tilde{T},\ \ldots$ .... Fuzzy Relation

Chapter 2

Fuzzy Logic

The nature of uncertainty in a problem is a very important point to select an appropriate method to represent vagueness. Probability is a theory for reasoning and decision making in environments with lack of certainty; on the other hand, it has been proved that FuSeTh is a mathematical way to represent, understand or approximate system behaviour when analytic functions or numerical relationships do not exist because the original system is involved with lack of data or human condition like those found in biological, medical, social, economic, political contexts [50].

In fact, from the theory arised from the Stone - Weierstrass approximation theorem [50], fuzzy systems can be successfully applied in contexts like highly complex systems whose behaviour are not well understood and, where a fast approximated solution is warranted. In this chapter we focus our attention in the understanding of the approximate reasoning and fuzzy inference in order to apply their concepts in the branch of fuzzy modeling.

2.1 Approximate Reasoning

Classical logic is a branch of mathematics which helps to make clear sentences and classical set theory provides a way to use these sentences which can be derived from objects that satisfy precise properties of membership [50]. However, in practical applications the propositions are usually assumed to be true or false and it is hard to precise how to use a sentence in a determined context since natural language and human communication are closely related. This is one of the reasons why multivalued logics are considered instead of the classical binary one. Fuzzy sets were introduced in order to be able to make calculus and computations with the interpretations of language. Roughly speaking, any function onto the unit interval can be used to approximate the membership in a set but it indeed characterizes a fuzzy set when some semantic description is associated to it. Thus, a method to formalize imprecise reasoning arises.

In first order logic, the simplest propositions are of the form subject-predicate. They are sentences for which it is possible to say if they are true or false without any doubt. Two simple propositions P and Q can be combined or modified by using one of the logical connectives, namely: conjunction ($\wedge$), disjunction ($\vee$), negation ($\neg$ or $\sim$), implication ($\rightarrow$) and equivalence ($\leftrightarrow$). By considering $T : U \rightarrow \{0, 1\}$ as a truth function, in the sense that $T(P) = 0$ means that the proposition P is false and $T(P) = 1$ when P is true, we have:

$$T(\neg P) = \begin{cases} 1 & T(P) = 0 \\ 0 & T(P) = 1. \end{cases} \tag{2.1}$$

$$T(P \wedge Q) = \min\{T(P), T(Q)\}. \tag{2.2}$$

$$T(P \vee Q) = \max\{T(P), T(Q)\}. \tag{2.3}$$

$$T(P \to Q) = \begin{cases} 0 & T(P) = 1 \text{ and } T(Q) = 0 \\ 1 & \text{Otherwise.} \end{cases} \tag{2.4}$$

$$T(P \leftrightarrow Q) = \begin{cases} 1 & T(P) = T(Q) \\ 0 & T(P) \neq T(Q). \end{cases} \tag{2.5}$$

From the previous ideas, we can find the **tautologies** which are compound propositions that are always true no matter the truth value of their components and the **contradictions** which are compound propositions which are always false, no matter the truth value of their components. Other two interesting compound propositions are **exclusive or** and the **exclusive nor**, the former is of interest because it arises in a lot of contexts that involve natural language or human reasoning; the last one, denoted as **XOR**, is the complement of the exclusive or. Tautologies are useful for making deductions and reasoning and two of the most common tautologies are the **Modus Ponens** and the **Modus Tollens** which are used in forward-chaining rule-based expert systems and in backward-chaining expert systems, respectively [50]; proofs by **Reductio Ad Absurdum** exploit the fact that $P \to Q$ is true if $P \wedge \neg Q$ is false. Another fact to focus in is the **paradox** because the restriction of classical propositional calculus to two-valued logic reduce it to half-truths but such conditions are not allowed since only $T(P) = 1$ or 0 is valid [50].

In real life, a variable is frequently described by words or sentences in natural language instead of numerical values. Linguistic statements tend to express subjective ideas and since natural language is fuzzy, because it involves vague and imprecise terms, linguistic variables become fundamental in most applications of fuzzy sets. Fuzzy logic seeks to form theoretical foundation for reasoning about imprecise propositions; such reasoning known as **approximate reasoning** is considered as an extension of classical propositional calculus which deals with partial truths and ilustrates the power of using fuzzy sets in the reasoning processes; also, this term is used to refer to logic inference based in a fuzzy version of the classic Modus Ponens [9, 50, 68, 94, 93].

The Modus Ponens deduction is used for making inferences in ruled-based systems but Buckley [60] comments that Modus Ponens is not useful in a fuzzy expert system as it does in classic logic. Therefore, approximate reasoning for fuzzy inference from IF-THEN rules like "if X is A, then Y is B" uses the **Generalized Modus Ponens** which is formulated as follows:

If X is A then Y is B.
From $X = A'$,
infer that $Y = B'$.

The elements A, A', B and B' are fuzzy sets. In fuzzy control, such elements are fuzzy numbers. Also, we can choose between many fuzzy implication operators but it depends on the application and this approximate reasoning may or not be consistent, that is, whenever $A' = A$ we can obtain $B' = B$ [9].

Natural language is a form of transmitting information for any problem or situation that requires to be reasoned or solved despite of its vagueness and ambiguity; therefore, the concept of linguistic variable plays an important role in representing imprecise human knowledge and it is a key notion to construct fuzzy systems in application areas like data analysis [48]. If a variable can assume linguistic terms as values, then it is called a **linguistic variable** [81] and it has associated a function which maps such a linguistic value into a fuzzy set in order to give it an "interpretation" for it. A linguistic value is a fuzzy set in nature!

In general, linguistic terms may be composed of a fuzzy predicate (also known as primary term [71] or atomic term [50]) which is a word or sentence whose meaning can contain ambiguity like young smart, small, slow, medium young, beautiful, so on; and **composite terms** which are collections of atomic terms built from atomic ones by using hedges (linguistic modifiers, adjectives or adverbs) to change the meaning of the predicative or by combining linguistic connectives (and, not, or) and their interpretation can be defined through set theoretic operations involving max, min or differences [50, 71, 81, 82, 83]; some examples of hedges are: very, likely, unlikely, extremely, almost, quite, more or less, mostly, few, all, usually, so on and, very slow, few dangerous, too young, beautiful and awesome, etc. are composite terms.

Based on both the Zadeh's definition [81, 82, 83] and the definition given by Bede [6], we present a the following mathematical definition for the concept linguistic variable.

Definition 2.1.1. *A **linguistic variable** x is a linguistic term characterized by the quintuple*

$$x := (x, U, T(x), M, \mathcal{G})$$

where,

- *x is the **name of the linguistic variable** (linguistic term).*
- *U is the **universe of discourse**, that is, where the characteristics of the variable can be defined (Also it is denoted by X).*
- *$T(x)$ is the set of **labels of linguistic values** of x.*
- *$M : T(x) \to \mathcal{F}(U)$ is a function called **semantic rule** (This function assigns to each label in $T(x)$ a mathematical object which will be called a fuzzy (sub)sets of U).*
- *$\mathcal{G}$ is named **syntactic grammar**. It produces linguistic values for x from composition of fuzzy sets and a certain type of functions called hedges.*

The following example clarifies the previous definition.

Example 2.1.1. *The linguistic term "Human age" can be transform in a linguistic variable if the quintuple $(Humanage, T, U, G, M)$ is considered. Here:*
$x =$ is Human age.
$U = [0, 120]$ is the universe of discourse for Human age.
$T(x) = \{$old, young, very old, very young, very very young, ...$\}$ is the set with the linguistic values taken into account for Human age.
$M : T(x) \longrightarrow \mathcal{F}(U)$ is a function such that $M(young) = Y$ and $M(old) = O$, where

$Y = (0, 18, 40)$ *and* $O = (35, 60, 80)$ *(These representation correspond to fuzzy triangular numbers and will be discussed later).*
G : *The syntax rules can be expressed as follows: If* $y \in T(x)$ *and very has associated the function* $h = x^2$*, then very young, with membership function* $h(M(y))$*, belongs to* $T(x)$*.*

Remark* 13.** *From Definition 2.1.1 and Example 2.1.1 we can see that a linguistic variable works as a kind of translator or interpreter that assigns to* ***linguistic terms fuzzy sets*. In addition, fuzzy sets offer a natural interaction between linguistic representations and numerical ones and provide a suitable tool for handling imprecise or ambiguous words; for instance, hedges may be interpreted as a composition between a given function and a basic membership function since they are formalized through functions* $h : [0, 1] \to [0, 1]$ *that satifies* $h(0) = 0$ *and* $h(1) = 1$ *[6, 21].*

There exist two types of predicates: imprecise and precise ones. A characteristic that linguistically distinguishes them is that in the fuzzy case once a linguistic term A and a hedge h are given, $h(P)$ is immediately understood but if P is precise, $h(P)$ might need a new definition to be understood. Example if P is *to be odd* in the set of natural numbers and h in *very* then $h(P)$ is *very odd* and we need to define what is understood by *very odd* [68].

Hedges are modifiers, adjectives, adverbs which change truth values of the fuzzy predicates and are fundamental to build fuzzy expert systems in the real world since they make writing rules easier and make programs more understandable for users and domain experts [9]. A **linguistic hedge** is mathematically defined as a unary operation on the unit interval which is biyective and satisfies that $h(0) = 0$ and $h(1) = 1$ in order to indicate that crisp predicates are not modified by the hedge [71]. There are several types of hedges but two are the most used in applications [60]:

1. Hedges that changes scalar numbers to fuzzy numbers with dispersion depending on the particular term used. The precise meaning of the hedge term will vary from one expert to another. The following table shows the most used hedges that create fuzzy numbers.

Hedge	**Spread ±% of central value at membership** 0.5
Nearly	5%
About	10%
Roughly	25%
Crudely	50%

Table 2.1. (Taken from [60]).

2. Hedges that modifies the truth value or membership function of a basic atomic term. They can be used to modify clauses in fuzzy propositions. They can be clasified in concentrations, dilations and intensification [80].

   (a) **Concentrations** reduce the degree of membership of all elements that are only "partly" in the set. Examples of concentrations are very (μ^2), extremely (μ^3) and plus ($\mu^{\frac{5}{4}}$).

(b) **Dilations** increase the degree of membership of all elements that are only "partly" in the set. Examples of dilations are slightly ($\mu^{\frac{1}{3}}$), somewhat ($\mu^{\frac{1}{2}}$) and minus ($\mu^{\frac{3}{4}}$).

(c) **Intensifications** are combinations of concentrations and dilations. Intensification increases the contrast between the element of the set that have more than half-membership and those which have less than half-membership [60] and they increase the degree of membership of all elements in the set with original membership values greater than 0.5 and they decrease the degree of membership of all elements in the set with original membership values less than 0.5. An example of intensifier is:

$$\begin{cases} 2\mu(x)^2 & 0 \leq \mu(x) \leq 0.5 \\ 1-2|1-\mu|^2 & 0.5 \leq \mu(x) \leq 1. \end{cases}$$

The use of a membership function gives the flexibility of an elastic meaning to a linguistic term. Composite terms can be formed from one or more combinations of atomic terms, logical connectives and linguistic hedges [60]. The implementation of linguistic hedges and logical connectives is manifested as function-theoretic operations on the values of the membership functions.

In literature, the following preferece table has been suggested for Standard Boolean Operations [60], parenthesis can be used to change the preceding order and ambiguities may be solved by the use of association to the right. For instance, "plus very minus very small" should be interpreted as

$$plus(very(minus(very(small)))).$$

Precedence	**Operation**
First	Hedge, not
Second	And
Third	Or

TABLE 2.2. Taken from [80].

In classical logic we have that a proposition is a sentence which is either true or false but not both. Nevertheless, if this sentence contains a fuzzy term, it becomes that proposition in a fuzzy one. For example the sentence "This dog smells stinky", the word "stinky" is a fuzzy term and makes a bit difficult the truth or falsity of the proposition [71]. A proposition with its truth value among 0 and 1 is called a **fuzzy proposition**. Note that a classical proposition is a special case of fuzzy proposition. There are two types of fuzzy propositions:

- Atomic propositions: They are sentences of the form "x is A" where x is a linguistic variable and A is a linguistic value of x.

- Compound fuzzy propositions: They are composition of atomic propositions through one or several logical connectives like and, or and not. In this type of fuzzy propositions may appear multiple linguistic variables and its truth value can be computed as follows [71]:

Let A, and B be linguistic values which are fuzzy sets on X and Y, respectively,

1. "x is NOT A", $N(A(x))$, N is the fuzzy negation.
2. "x is A AND y is B", $T(A(x), B(x))$, T is a $T-$norm.
3. "x is A OR y is B", $S(A(x), B(x))$, S is a $T-$conorm.

In the book written by Klir and Yuan [33] we can find four types of fuzzy propositions which are of the form:

- "x is A". Example, $Temperature$ is Low.
- "x is $h(A)$", where h is a hedge or modifier. Example, $Temperature$ is $Very\ low$.
- "IF x is A THEN y is B" (Generalized Modus Ponens). Example, IF $Temperature$ is Low THEN $Lettuce$ are $Fresh$.
- "IF x is A THEN y is $h(B)$" Example, IF $Temperature$ is Low THEN $Lettuce$ are $QuiteFresh$.

Although the number of terms in a linguistic variable can be large, usually it is considered between five and nine since, in a lot of cases, less than five is poor and more than nine is excessive. Anyway in a good number of applications there only appear three terms P, its opposite aP and MP, where MP is equivalent to NOT P AND NOT aP, i.e., MP satisfies [68]:

$$\mu_{MP}(x) = \min\{\neg P(x),\ \neg aP(x)\}.$$

To define aP a symmetry on X, $A : X \to X$, is needed. That is A satisfies:

(i) If $x \preceq y$ then $A(y) \leq A(x)$.

(ii) $A \circ A = id_X$.

Note that $aP(x) := P \circ A(x) = P(A(x))$ for all $x \in X$. Thus, $a(aP)(x) = aP(A(x)) = P(A(A((x))) = P(x)$. When modeling $aP = P \circ A$ and $\neg P = N \circ P$ with a symmetry A on X and a strong negation N in $[0, 1]$, it results the **condition of coherence** [68]:

$$P \circ A \leq N \circ P.$$

Thus, if A is know, N should be chosen to satisfy this coherence condition and vice versa.

To end this section up, we will introduce the concept of IF-THEN rule from a set theorical approach [68]. A **logic** is a triplet $(X, \mathcal{A}, C)$ where X is a crisp set, $\mathcal{A} \subseteq \wp(X)$ and $C : \mathcal{A} \to \mathcal{A}$ is a function which is:

⋆ **Extensive**: For all $P \in \mathcal{A}$, $P \in \mathcal{A}$ implies $P \subseteq C(P)$.

⋆ **Monotonic**: For all $P, Q \in \mathcal{A}$, $P \subseteq Q$ implies $C(P) \subseteq C(Q)$.

⋆ **Closure**: For all $P \in \mathcal{A}$, $C(C(P)) = C(P)$, i.e., $C^2 = C$.

The following examples were taken from Trillas's book [68].

Example 2.1.2. *Let* $(X, \cdot, +, 0, 1)$ *be a finite lattice.* $\mathcal{A} = \wp_0(X) := \{P = \{p_1, \dots, p_n\} \subseteq X \mid p_\wedge := p_1 p_2 \cdots p_n \neq 0\}$ *and Cons* : $\wp_0(X) \to \wp_0(X)$ *defined by* $Cons(P) = \{q \in X \mid p_\wedge \preceq q\}$ *where the partial order* $\preceq$ *is given by* $x \preceq y \Leftrightarrow x \cdot y = x$. *Thus, Cons is a consequence operator.*

Example 2.1.3. *Let us take the set* $[0, 1]^X$ *endowed with a fuzzy* $T-$*norm*

$$\mu_\wedge = \mu_1 \cdots \mu_n = T \circ (\mu_1 \times \cdots \times \mu_n),$$

the partial order $\mu \preceq \sigma \Leftrightarrow \mu(x) \leq \sigma(x)$ *for all* $x \in X$ *and the empty set* $\mu_0 = \mu(\emptyset)$. *Consider the set*

$$\mathbb{P}_0\left([0, 1]^X\right) := \left\{P = \{\mu_1, \cdots, \mu_n\} \subseteq [0, 1]^X \mid \mu_\wedge \neq \mu_0\right\}.$$

The definition

$$Cons(P) = \{\sigma \in [0, 1]^X \mid \mu_\wedge \preceq \sigma\}.$$

The process to pass from a set of premises $P = \{\mu_1, \cdots, \mu_n\}$ to a "conclusion" σ is known as **conclusive reasoning** and it is symbolized by $P \vdash \sigma$. We say that it is a **deductive reasoning** when there exists an operator of consequences C sucha that $P \vdash \sigma$ is equivalent to $\sigma \in C(P)$ but, not all conclusive reasoning are deductive; additionally, the most common are not deductive but **conjectural**, i.e., there exists an operator of consequences C such that $P \vdash \sigma \Leftrightarrow N \circ \sigma \notin C(P)$ for some strong negation N [68].

If we consider the operator $Cons(P) = \{\sigma \mid \mu_\wedge \preceq \sigma\}$ provided $\mu_\wedge$ (like in Example 2.1.3) then, the reasoning is named a **not self-contradictory** (i.e., $\mu_\wedge \npreceq N \circ \mu_\wedge$). Notice that the par (T, N) is necessary to the definition of $Conj(P)$, where

$$Conj(P) := Cons(P) \cup Hyp(P) \cup Sp(P),$$

where the set of hypothesis of P is

$$Hyp(P) := \{\sigma \in Conj(P) \mid \mu_0 < \sigma < \mu_\wedge\}$$

and the set of speculations is

$$Sp(P) := \{\sigma \in Conj(P) \mid \mu_\wedge \text{ is not comparable under } \preceq \text{ with } \sigma\}.$$

This union is a disjoint union; speculations are the conjetures for which it is neither $\mu_\wedge \preceq \sigma$ nor $\sigma < \mu_\wedge$. In addition [68],

- $P \vdash \sigma$ is a **guessing** if $\sigma \in Conj(P)$.
- $P \vdash \sigma$ is a **deduction** if $\sigma \in Cons(P)$.
- $P \vdash \sigma$ is a **abduction** if $\sigma \in Hyp(P)$.
- $P \vdash \sigma$ is a **speculation** if $\sigma \in Sp(P)$.

Carrying on this treatment, a **conditional statement** is a sentence of the form "IF p THEN q", where p and q are propositions, and it is denoted $p \to q$. We are talking about them because imprecise conditions (rules) are useful in common life and in technology [68]. That is why is necessary to give a representation for them.

In order to represent a conditional statement like "IF x is P THEN y is Q", it is required to translate it in fuzzy terms with adequate fuzzy sets. In fuzzy logic, it is always be done by using a function $J : [0,1] \times [0,1] \to [0,1]$ such that for all $x \in X$ and $y \in Y$:

$$(P \to Q)(x,y) \equiv P(x) \to Q(y) := J(P(x), Q(y)) \in [0,1].$$

This function depends on the "meaning" of the conditional statement and requires to look at what happens in general [68] in order to take premises "x is P" and "IF x is P THEN y is Q" and conclude "y is Q". Formally, it should exist a continuous $T-$norm T_0 such that for all $x \in X$ and $y \in Y$:

$$T_0(P(x), J(P(x), Q(y))) \leq Q(y).$$

This inequatilty is called **Modus Ponens Inequality (MP-Inequality)** and J will be named T_0-**conditional** function [68].

***Remark* 14.** *The function J is called $Q-$**conditional**, if it can be expressed as $J(a,b) = S(c(a), T(a,b))$ where S is a continuous $T-$conorm, T is a continuous $T-$norm and c is a strong negation; if $J(a,b) = S(c(a), b)$ for all $a, b \in [0,1]$ and some $T-$conorm S then J is an $S-$**implication**.*

These rules are also known as **IF-THEN rules** because they are commonly represented by using fuzzy implications operators [71]. Some operation that represent the fuzzy IF-THEN rules are:

Dienes-Rescher	$\max\{1 - P(x), Q(y)\}.$	(2.6)
Lukasiewicz	$\min\{1, 1 - P(x) + Q(y)\}.$	(2.7)
Gödel	$\begin{cases} 1 & P(x) \leq Q(y) \\ Q(y) & Otherwise. \end{cases}$	(2.8)
Mamdani	$\min\{P(x), Q(y)\}.$	(2.9)
Zadeh	$\max\{\min\{P(x), Q(y)\}, 1 - P(x)\}.$	(2.10)

2.2 Fuzzy Inference

In classic logic, there are three important fuzzy inference rules [6], namely:

- Modus Ponens

$$\frac{\begin{array}{l} A \to B \\ A \end{array}}{B}$$

- Modus Tollens

$$\frac{\begin{array}{r} A \to B \\ \neg B \end{array}}{\neg A}$$

- Hypotetical Sylogism

$$\frac{\begin{array}{c} A \to B \\ B \to C \end{array}}{A \to C}$$

In practice, we frequently make imprecise reasoning and to be able to model expert opinions or common sense knowledge we must think in rules which deal with them and linguistic terms: **Fuzzy rules** (also known as **Fuzzy IF-THEN rules**).

A **fuzzy rule** is a triplet $(A, B, \mathcal{R})$ that consist of an antecedent $A \in \mathcal{F}(X)$, a consequence $B \in \mathcal{F}(Y)$ which are linguistic variables linked through a fuzzy relation $\mathcal{R} \in \mathcal{F}(X \times Y)$. for example from "IF $x \in A$ THEN $y \in B$", we can define it as a fuzzy relation as follows:

(i) **Mamadani Rule**:
$$\mathcal{R}_M(x, y) := A(x) \wedge B(y).$$

(ii) **Larsen Rule**:
$$\mathcal{R}_L(x, y) := A(x)B(y).$$

(iii) $T-$**norm Rule**: Let T be a $T-$norm,
$$\mathcal{R}_T(x, y) := T(A(x), B(y)).$$

(iv) **Gödel Rule**:
$$\mathcal{R}_G(x, y) := A(x) \to B(y).$$

***Remark* 15.** *In practical problems, having not a unique way to interpret a fuzzy rule is an advantage because we have more flexibility to select the method that best fits our problem out [6].*

In fuzzy logic, inference rules are generalized as follows [71]:

- Generalized Modus Ponens (GMP)
$$\begin{array}{ccc} A & \to & B \\ A' & & \\ \hline & & B' \end{array}$$

- Generalized Modus Tollens (GMT)
$$\begin{array}{ccc} A & \to & B \\ & & B' \\ \hline A' & & \end{array}$$

- Generalized Hypotetical Sylogism (GHS)
$$\begin{array}{ccc} A & \to & B \\ B' & \to & C \\ \hline A & \to & C' \end{array}$$

To calculate the fuzzy set B' in the GMP let $A \in \mathcal{F}(X)$ and $B \in \mathcal{F}(Y)$. The proposition "IF $x \in A$ THEN $y \in B$" is expressed as a fuzzy relation $\mathcal{R} \in \mathcal{F}(X \times Y)$. Then, finding the $T-$composition of A' and $\mathcal{R}$ in order to obtain the conclusion B', i.e., if T is a $T - norm$, for all $y \in Y$:
$$B'(y) = \max_{x \in X} T(B'(y), \mathcal{R}(x, y)).$$

For the GMT, the mathematical representation of reasoning is similar. Firstly, "IF $x \in A$ THEN $y \in B$" is expressed as a fuzzy relation $\mathcal{R} \in \mathcal{F}(X \times Y)$. The B' and $\mathcal{R}$ are combined by $T-$composition to obtain the conclusion A' as follows:
$$A'(x) = \max_{x \in X} T(B'(y), \mathcal{R}(x, y)).$$

To end, let us turn to the GHS. The implications "IF $x \in A$ THEN $y \in B$" and "IF $y \in B'$ THEN $z \in C$" are expressed by fuzzy relations $\mathcal{R}_1$ and $\mathcal{R}_2$; then, the fuzzy proposition "IF $x \in A$ THEN $z \in C'$" is represented by the composition $\mathcal{R} := \mathcal{R}_2 \circ \mathcal{R}_1$.

A property which is required for a fuzzy system is the **interpolation property**. It says that when the input of the system coincides with the antecedent then the output has to be coincident with the consequence, i.e., if $A' = A$ then $B' = B$ [6].

Proposition 2.2.1. *The Mamdani inference with $T-$norm rule*

$$B'(y) = \max_{x \in X} \min\{A'(x), A(x) \wedge B(y)\}$$

*such that there exists an $x_0 \in X$ with $A(x_0) = 1$, satisfies the **interpolation property**, i.e., IF $A' = A$ THEN $B' = B$.*

Proof. From the properties of $T-$norm we have that $A(x) \wedge B(y) \leq A(x)$. Thus,

$$\begin{aligned} B'(y) &= \max_{x \in X} \min\{A'(x), A(x) \wedge B(y)\} \\ &= \max_{x \in X} \min\{A(x), A(x) \wedge B(y)\} \\ &= \max_{x \in X} A(x) \wedge B(y) \\ &= \left(\max_{x \in X} A(x)\right) \wedge B(y) \\ &= 1 \wedge B(y) \\ &= B(y). \end{aligned}$$

□

To end this section up, we present the three types of rule-based inference found in the Pedrycz' book [48]:

- ⋆ **Monotonic**: The consequent truth values may increase but not decrease by any new information.
- ⋆ **Non-Monotonic**: The consequent truth values may increase or decrease; the new information supplied is more reliable than any existing information.
- ⋆ **Downward Monotonic**: The consequent truth values may decrease but not increase. This type of inference is useful when modifying multivalued data, necessary when defuzzifying.

For any linguistic variable, we can use three different general forms to create canonical rules [62]. They are:

- ⋆ **Assigment statement**: Here the variable has assigned a value which can be a linguistic term; this values is restricted to a especific equalty. For example, $y = high$; $climate = warm$; $a = 5$.
- ⋆ **Conditional statement**: They are fuzzy conditional statements. For example, "IF $Mark > 50$ points THEN *pass*.
- ⋆ **Unconditional statement**: There is no specific condition to be satisfied. For example, *stop*; *Push the valve.*

The process of obtaining a conclusion from all consequences in the fuzzy rules is known as **aggregation** (of the rule). The two methods for determining them are: **Conjuctive** and **Disjuctive Systems**. The former are connected by AND, the latter with OR connectives.

Fuzzy inference is a process where a conclusion is obtained from a given input. The basic rule for a inference system is the compositional rule of inference which is based on the classical rule of Modus Ponens [62].

Given for a fuzzy rule $\mathcal{R} \in \mathcal{F}(X \times Y)$, the **compositional rule of inference** is a function

$$F : \mathcal{F}(X) \rightarrow \mathcal{F}(Y) \tag{2.11}$$

$$A' \mapsto F(A') = A' \circ \mathcal{R} \tag{2.12}$$

where $\circ : \mathcal{F}(X) \times \mathcal{F}(X \times Y) \rightarrow \mathcal{F}(Y)$ is a composition of fuzzy relations. The relation $\mathcal{R}(x, y)$ is a fuzzy relation which is used for interpretating the fuzzy rule in the premise [6].

2.3 Fuzzy Modeling

Why do we need fuzzy systems when we have a crisp relationship, crisp input and also a crisp output for a classical system? Because it is often in practical applications and systems to find uncertainty due to relationships that are unknown or partially known to us and whose analysis could require the understanding and implementation of common sense knowledge. Fuzzy systems can fill the gaps in and approximate any desired output with arbitrary precision by combining approximation properties of the fuzzy systems with the ability to model linguistic expressions [68]. The subjectivity that exists in fuzy modeling is "a blessing rather that a curse" [50] and fuzzy systems are tools which provide a means of sharing, communicating and transferring this human subjective knowledge of systems and processes. The results in the form of fuzzy sets are benefical for the interpretation purposes of fuzzy modeling [48]. Thus, we find a very interesting field to study which can be efficient to applications.

Fuzzy inference systems (FIS), also known as fuzzy rule-based systems, fuzzy models or fuzzy expert systems, are mainly based on some concepts of the fuzzy set theory like linguistic variables, fuzzy rules and fuzzy inference methods. The linguistic variables allow us to interpret linguistic expressions in terms of fuzzy mathematics, the fuzzy rules (like fuzzy IF-THEN rules) are a set of laws that associate inputs with outputs and, the **inference mechanism** is a tool to model the process of approximate reasoning through fuzzy rules. A **fuzzy rule base** is a finite collection of fuzzy rules whose elements are naturally translated into fuzzy relations by composition of the type $\max - \min$ or $\min \rightarrow$ [6].

A fuzzy system consists of a fuzzification interface, a rule base, a database, a decision making unit (DMU) and a defuzzification interface. In the FIGURE 2.1 it is illustrated the basic structure of a fuzzy inference system where the rule base contains IF-THEN rules, the database defines the membership functions of the fuzzy sets used in the fuzzy rules, The DMU performs the inference operations on the rules and the fuzzification interface transforms crisp inputs into fuzzy sets in order to construct the processing core of the model and so they become more informative and comprehensible than a single numeric quantity. The defuzzification interface transforms the fuzzy results into crisp outputs.

The steps of fuzzy reasoning performed by a fuzzy inference system are [62]:

1. **Fuzzification**: Obtain the membership values of each linguistic label.

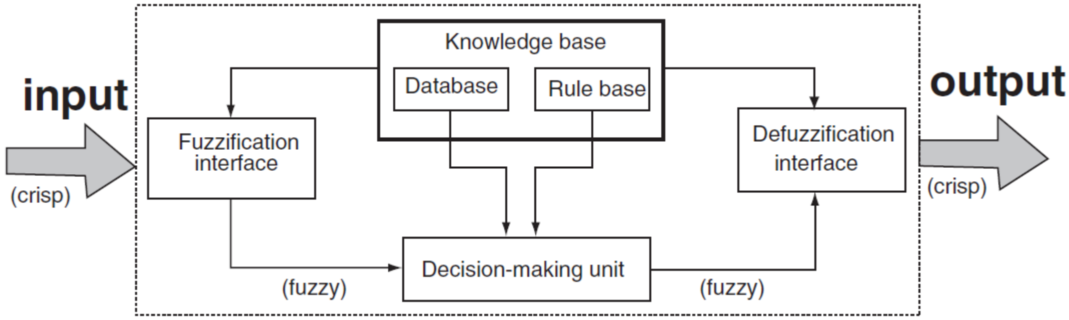

FIGURE 2.1. [62].

2. **To get the "Firing Strength"** of each rule: By using $T-$norms combine the membership values and the premise in order to obtain the degree that satisfies a fuzzy rule.

3. **Generate the qualified consequent**.

4. **Defuzzification**: Combine (Aggregate) the consequent of each input to produce a crisp output.

The working of a Fuzzy Inference System can be summarizes as follows: First the crisp input is transformed in a fuzzy one by some fuzzification method then, the rule base and the data base are defined in order to do fuzzy reasonings. Finally the fuzzy output is converted into a crisp one by using a deffuzzification method [48]. Now we present some ideas about categories of fuzzy models. In literature, it is possible to find different architectures in the area of fuzzy modeling and some of them are:

1. **Tabular Fuzzy Model (TFM)**. It is formed by tables of relationships between the variables of the system. This type of model produce a compact suite of relationships represented at the level of information granules. The evident advantage of them resides with their evident readability.

	B_1	B_2	B_3	B_4	B_5
A_1					
A_2			C_3		
A_3					C_1

FIGURE 2.2. [48].

2. **Rule-based Systems**. They are composed of a family of fuzzy conditionals or IF-THEN statements (called rules) where fuzzy sets occur in their conditions and conclusions. Thys type of model is very important, that is why it is exposed in other place of this text.

3. **Fuzzy Relational Models (FRMs)**. They are examples of ideas whose computing let us encounter a wealth of architectures due to the variety of composition operators. The FIGURE 2.3 shows a general taxonomy of fuzzy relational equations (modeling structures) presented with respect to the combination of composition operators used in their realization.

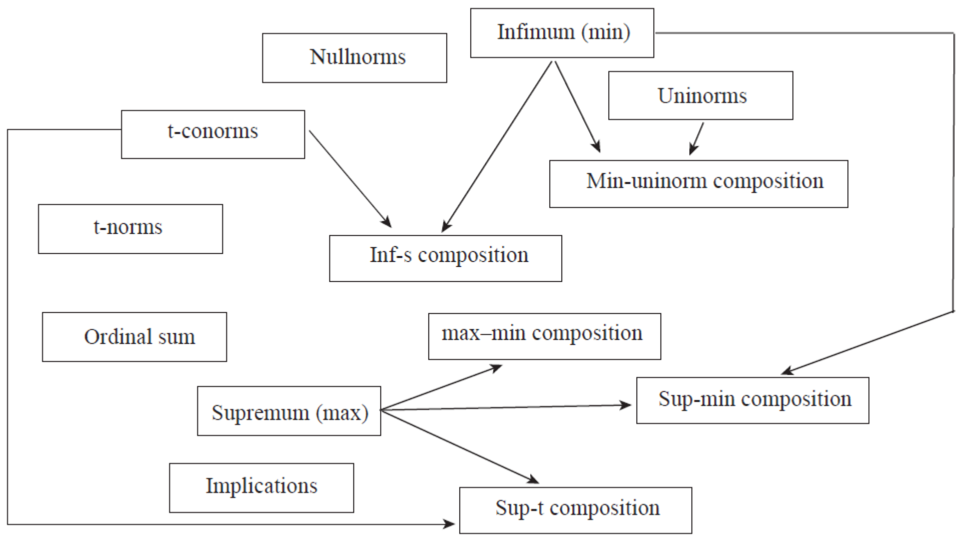

FIGURE 2.3. [48].

4. **Fuzzy Decision Trees (FDTs)**. They benefit from the inductive learning approach that underpins their construction, to aid in the classification of objects based on their values over different attribute. Their construction in a fuzzy environment allows for the potential critical effects of imprecision to be mitigated through the decision rules defined.

5. **Fuzzy Neural Networks (FNNs)**. It is also know as neuro-fuzzy system and it is a learning machine that finds the parameters of a fuzzy system through approximation techniques from neural networks.

Fuzzy inference systems are the most important tool based on fuzzy set theory. They are built by domain experts and used in automatic control, decision analysis and various other expert systems [50]. To finish this part of the document devoted to Fuzzy Logic, we talk about the most important types of **fuzzy inference methods** (FIM): The Mamdani FIM and the Takagi-Sugeno-Kang Method (TSK Method). The main difference between them lies in the consequences of fuzzy rules. Mamdani fuzzy systems use fuzzy sets as rule consequent whereas TST fuzzy systems employ linear functions of input variables as rule consequent. Among their advantages Sugeno's method is computationally efficient because its well suited mathematical analysis makes it to work well with linear, optimization and adaptive techniques; On the other hand, Mamdani is intuitive, with a widespread acceptance and it is well suited to human inputs.

The result of Sugeno Reasoning is an exact number. In Sugeno reasoning the consequence is deterministic. The input data is a finite set $\{(x_i, y_i)\}_{i=0}^{n}$. The rules are often of the form IF x is x_i THEN $y = y_i$ where the product is used for THEN-implication. Sugeno reasoning allows us to use functions of input x not only constants. Thus, the rules would like of the form IF x is x_i THEN $y = f_i(x)$ where each function f_i can be different nonlinear mapping.

2.3.1 Mamdani Fuzzy Inference System

This method was proposed by Mamdani in 1975 and it was based on a Zade's paper [80]. To compute the output a fuzzy inference system with Mamdani method, the following steps has to be followed [62]:

1. Determine a set of fuzzy rules.
2. Fuzzify the inputs by using some fuzzification method.
3. Combine the fuzzy inputs according to the fuzzy rules.
4. Find the consequence of the rule.
5. Combine the consequence.
6. Defuzzify the output by using Center of Mass, Mean of Maximum.

***Remark* 16.** *The Mamdani method has several variations. There are different $T-norms$ to use for the connectives of the antecedents, different aggregation operators for the rules and numerous defuzzification methods that could be used [50].*

2.3.2 Sugeno Fuzzy Inference System

It was proposed by Takagi and Sugeno in 1985 to generate fuzzy rules from a given input-output data set [50]. A typical rule in a Sugeno model has the form

$$\text{IF } x \in A \text{ AND } y \in B \text{ THEN } z = f(x, y)$$

where f is a crisp set function usually it is a polynomial function in the outputs x and y, when $f(x, y)$ is a constant, the inference system is called $zero - order$ Sugeno model, which is a special case of the Mamdani system in which each rule's consequent is especified as a fuzzy singleton. In a Sugeno model, each rule has a crisp output, given by a function.

Input the crisp value x_0.
Compute the firing strenght of each fuzzy rule

$$\alpha_i := A_i(x_0).$$

Compute the rules ouputs

$$z_i := a_i x_0 + b_i.$$

The system output is

$$z_0 := \frac{\sum_{i=1}^{n} \alpha_i z_i}{\sum_{i=1}^{n} \alpha_i}.$$

Other case:

Input the crisp value x_0 and y_0.
Compute the firing strenght of each fuzzy rule

$$\alpha_i := A_i(x_0) \wedge B_i(x_0).$$

Compute the rules ouputs

$$z_i := a_i x_0 + b_i y_0 + c_i.$$

The system output is

$$z_0 := \frac{\sum_{i=1}^{n} \alpha_i z_i}{\sum_{i=1}^{n} \alpha_i}.$$

***Remark* 17.** *Sugeno systems that use constants or linear functions in the consequence are parametrized maps. Hence, it is possible to use optimization techniques to find the best parameters to fit data. The structure is similar although the expression depends on the choice of the $T-$norm and the defuzzification.*

2.3.3 Mamdani vs Sugeno

In data mining clustering techniques are employed through a large volume of data to acquire relevant and significant data in pattern recognition; i.e., one of the goals of data mining is data pattern recognition. According to Moreno et al, the subtractive clustering technique within the Takagi-Sugeno-Kang FIS method, applied on specific databases, presents a better performance than any other technique because it best optimizes the objective function and the value of the root mean square error is minimized enormously; although, in terms of handling linguistic variables, Mamdani is a better choice and its respective error values are very good too [43].

Saepullah et alia analysed and compared Mamdani, Sugeno and Tsukamoto methods on FIS to find a best method in terms of reduction in electrical energy consumption of air conditioner by using Room Temperature and Humidity as input variables and Compressor speed as output variable. He concluded that although the results were varied, not only the best method in terms of reduction of electrical energy consumption is Tsukamoto method but the Mamdani method is better than the Sugeno one [52].

Zaher et alia presented a comparison between FIS for prediction with application to prices of fund (in Egypt) because he says that the Fuzzy logic is a relatively modern method closer to human thinking and natural language. He remarked that, in this context, the Sugeno FIS is more flexible because it allows more parameters in the output, computationally is more effective and its structure for the rule outputs is more convenient for functional analysis than a Mamdani FIS [24].

In the comparison of Mamdani and Sugeno FIS for a satellite image classification, made by Salman and Seno, four types of fuzzy membership function generation methods that generate fuzzy IF-THEN rules directly from training data were considered. The research showed that, for this case study, Mamdani FIS does not only works better but also shows a more robust structure in the presence of noisy input data while a Sugeno FIS often require a cumbersome and time consuming task due to it implies th fact of dividing a multiple input multiple output system (MIMO)into as many multiple input single output systems (MISO) as the number of output variables [53].

To end, it can be concluded that Mamdani-type FIS and Sugeno-type FIS perform quite similar, but Sugeno-type FIS allows the evaluation of Breast cancer risk to work at its full capacity with smooth operational performance. Sugeno-type FIS has also an advantage that it can integrated with neural networks and genetic algorithm or other optimization techniques so that the system can adapt to system characteristic efficiently [59].

Chapter 3

Fuzzy Measure Theory

As we have mentioned before, FuSeTh is an appropiate tool which helps us to represent and communicate information, opinions and human decision making. The fuzzy measures were introduced by Michio Sugeno in 1974 (see [64]) in order to measure uncertainty that depends mainly on human subjectivity or by hesitation [3] but, in the literature, this term does not imply to apply a measure to fuzzy sets neither it deals with fuzzy measures in the sense of the FuSeTh but to non-negative monotonous functions called generalized measures [72].

Fuzzy measures have started to be accepted by researchers and, for the scientific community it has become an interesting branch of mathematics where we can find things to explore and applicate in disciplines like artificial intelligence, psychology, decision theory, economics, image processing, etc [3]. Fuzzy measure theory is a field of research which lies at the intersection of measure theory, aggregation functions theory, theoretical computer science among others [28]. In this part of the document, we present some of the most relevant concepts and results about fuzzy measures developed to this day in order to support the theory of possibility which is considered as a complement of the probability theory to measure uncertainty.

3.1 Why fuzzy measures?

The Classic Measure Theory (CMT) is one of the biggest areas of the mathematical analysis and the notion of measure let us generalize concepts like length, area, volumen, even it has been applied in different contexts including the modeling of uncertainty based on randomness [3]. The main feature of the classic measures is the $\sigma-$additivity which is useful in several applications of economics and statistics; however, there are contexts related to artificial intelligence, neural networks, decision making where the property of additivity, which characterize the classic measures, limits the power of this theory because it can be inflexible in the sense of it implicitly assumes that there are not any interaction between the elements of a set and their measurements [3, 77].

There are various types of measures: Counting measure, Lebesgue measure, Monotone measure, Probability measure, so on and, one of the most known applications of the CMT is in the probability theory which allows us to measure the change of occurrence of an event with a number [15]. Nevertheless and in spite of the probability measure is useful when uncertainty is based on randomness, the no random uncertainty of a real situation like uncertainty by vagueness and uncertainty by hesitation which are subjective makes probability be unable to give an accurate measurament. So, a fuzzy measure μ arises as a flexible extension of probability measure in the following sense: The additivity property is weakened by considering a condition of monotonicity, namely, If $A, B \in \mathcal{F}$ and $A \subseteq B$ then $\mu(A) \leq \mu(B)$. We will refer this extension as non-additive measure like in [39]. Sugeno was the first in proposing a definition for fuzzy measure but nowdays this concept is associated to generalized measures; when these measures are defined on fuzzy sets they are called fuzzified monotonous measures [72]. In this chapter we study the theory proposed by Sugeno.

3.2 Fuzzy measures

In order to define a fuzzy measure, it is necessary to consider a **measurable space** $(X, \mathcal{F})$ where X is a crisp set and $\mathcal{F} \subseteq \wp(X)$ is a $\sigma-$**algebra**. This family of subsets of X, according with Bartle [5], must satisfy:

1. $\emptyset \in \mathcal{F}$;
2. $A^c \in \mathcal{F}$ whenever $A \in \mathcal{F}$; and
3. Whether $\{A_n\}_{n\in\mathbb{Z}^+} \subseteq \mathcal{F}$ then $\bigcup\limits_{n\in\mathbb{Z}^+} A_n \in \mathcal{F}$.

Fuzzy measures, also termed non-additive measures [18], capacities [13] or monotone measures [72], are a generalization of classical ones because they are characterized by making $\sigma-$additivity more flexible. This property is replaced by a weaker one known as monotony. Thus, the following definition for a fuzzy measure is presented.

Definition 3.2.1. *Let $(X, \mathcal{F})$ be a measurable space. A **fuzzy measure** is a function $\mu : \mathcal{F} \to [0, \infty)$ that accomplishes:*

(FM1) $\mu(\emptyset) = 0$. ***(Boundary condition).***

(FM2) *For all $A, B \in \mathcal{F}$, if $A \subseteq B$ then $\mu(A) \leq \mu(B)$.* ***(Monotony).***

With this definition, **every crisp measure can be treated as a fuzzy measure**. When $\mu(X) = 1$, the fuzzy measure is called **normalized fuzzy measure**. If the condition **(FM1)** is skipped, the function μ is termed **non-monotonic fuzzy measure** (also, it is known as **game** and denoted by the Greek letter ν [28]). Sometimes it is desirable that μ can satisfy one or both of the following conditions:

(LC) If $\{B_n\}_{n\in\mathbb{Z}^+} \subseteq \mathcal{F}$ and $B_1 \subseteq B_2 \subseteq \cdots$ then

$$\mu\left(\bigcup_{n\in\mathbb{Z}^+} B_n\right) = \lim_{n\to\infty} \mu(B_n). \text{ (\textbf{Lower Continuity})}$$

(UC) If $\{B_n\}_{n\in\mathbb{Z}^+} \subseteq \mathcal{F}$, $B_1 \supseteq B_2 \supseteq \cdots$ and $\mu(B_1) < \infty$ then

$$\mu\left(\bigcap_{n\in\mathbb{Z}^+} B_n\right) = \lim_{n\to\infty} \mu(B_n). \text{ (\textbf{Upper Continuity})}$$

A fuzzy set that satisfies **(FM1), (FM2), (LC)** and **(UC)** is termed **Sugeno's Fuzzy Measure (SuFuMe)** (also it is known as **Continuous Fuzzy Measure**). The SuFuMe is useful fuzzy measure to characterize uncertainty but be aware of saying that fuzzy measures are SuFuMes.

Example 3.2.1. *(**Fuzzy measure that is not a SuFuMe**) Let $X = \mathbb{R}$, $\mathcal{F} = \wp(X)$ and $\mu : \wp(X) \to [0, 1]$ be a function defined by:*

$$\mu(A) = \begin{cases} 0 & A = \emptyset, \\ 1 & A \neq \emptyset. \end{cases}$$

*Note that, because of its definition, μ satisfies **(MF1)** and **(MF2)**. Hence, μ is a fuzzy measure but not a SuFuMe because if we consider the sequence defined by $B_n = \left(0, \frac{1}{n^2}\right)$ then $\mu\left(\bigcap_{n\in\mathbb{Z}^+} B_n\right) = \mu(\emptyset) = 0$ but, $\lim_{n\to\infty} \mu(B_n) = 1$.*

*Since $B_n \neq \emptyset$ for all $n \in \mathbb{Z}^+$, then μ does not hold the condition **(UC)**.*

Other definitions related to fuzzy measures are presented below.

Definition 3.2.2. *A fuzzy measure μ is said to be:*

- ***Convex (Supermodular)*** *whether for all $A, B \in \mathcal{F}$,*

$$\mu(A) + \mu(B) leq \mu(A \cup B) + \mu(A \cap B).$$

- ***Concave (Submodular)*** *whether for all $A, B \in \mathcal{F}$,*

$$\mu(A \cup B) + \mu(A \cap B) \leq \mu(A) + \mu(B).$$

- ***Maxitive*** *whether for all $A, B \in \mathcal{F}$,*

$$\mu(A \cup B) = \max\{\mu(A), \mu(B)\}.$$

- ***Minitive*** *whether for all $A, B \in \mathcal{F}$,*

$$\mu(A \cap B) = \min\{\mu(A), \mu(B)\}.$$

- ***Symmetric*** *when for all $A, B \in \mathcal{F}$, $|A| = |B|$ implies $\mu(A) = \mu(B)$.*

When $|X| = n$, we have a finite fuzzy measure and it is said to be:

- $k-$***alternating*** *(*$2 \leq k \leq n$*) whether for any family of* k *sets* $A_1, A_2, \ldots, A_k \in \mathcal{F}$,

$$\mu\left(\bigcap_{i=1}^{k} A_i\right) \geq \sum_{\substack{I \subseteq \{1,2,\ldots,k\} \\ I \neq \emptyset}} (-1)^{|I|+1} \mu\left(\bigcup_{i \in I} A_i\right).$$

- $k-$***monotonic*** *(*$2 \leq k \leq n$*) if for any family of* k *sets* $A_1, A_2, \ldots, A_k \in \mathcal{F}$,

$$\mu\left(\bigcup_{i=1}^{k} A_i\right) \geq \sum_{\substack{I \subseteq \{1,2,\ldots,k\} \\ I \neq \emptyset}} (-1)^{|I|+1} \mu\left(\bigcap_{i \in I} A_i\right).$$

- ***Totally monotone*** *when for all* $2 \leq k \leq n$*,* μ *is* $k-$*monotonic.*

Finite fuzzy measures are important since most practical applications involve finite sets [3].In addition, **When** X **is a finite set, i.e.** $|X| = n$**, any fuzzy measure is a SuFuMe** since any monotonous sequence of its subsets becomes in a stationary one and therefore the conditions **(UC)** and **(LC)** are immediately satisfied and some additional properties for μ can be found. The next proposition gives two properties that every fuzzy measure satisfies.

Theorem 3.2.1. *Let* μ *be a fuzzy measure on* $(X, \mathcal{F})$*, then*

1. *For all* $A \in \mathcal{F}$*,* $\mu(A) \geq 0$.
2. *For all* $A, B \in \mathcal{F}$*,*

$$\mu(A \cap B) \leq \min\{\mu(A), \mu(B)\} \leq \max\{\mu(A), \mu(B)\} \leq \mu(A \cup B).$$

Proof. Let μ be a fuzzy measure on $(X, \mathcal{F})$.

1. Let $A \in \mathcal{F}$. Since $\emptyset \subseteq A$ and μ is monotonous then $0 = \mu(\emptyset) \leq \mu(A)$.
2. Let $A, B \in \mathcal{F}$. Since μ is monotonous and

$$A \cap B \subseteq A \subseteq A \cup B \qquad \text{and} \qquad A \cap B \subseteq B \subseteq A \cup B,$$

then, it is obtained that

$$\mu(A \cap B) \leq \mu(A) \leq \mu(A \cup B) \qquad \text{and} \qquad \mu(A \cap B) \leq \mu(B) \leq \mu(A \cup B).$$

Hence,

$$\mu(A \cap B) \leq \min\{\mu(A), \mu(B)\} \leq \max\{\mu(A), \mu(B)\} \leq \mu(A \cup B).$$

□

***Remark* 18.** *The understanding of the fuzzy measures properties is useful in applications because when a fuzzy measure is used to define a function such as the Sugeno integral or Choquet integral (they will be define further) these properties will be crucial in understanding the function's behavior [67]; for instance, the Choquet integral with respect to an* *additive fuzzy measure reduces to the Lebesgue integral [3], submodular fuzzy measures result in convex functions while supermodular fuzzy measures result in concave functions[28].*

A method to build examples of fuzzy measures from a crisp one is given by the following result.

Theorem 3.2.2. *Let α be a crisp measure and $T : [0, \infty) \to [0, \infty)$ a real and monotonously increasing function such that $T(0) = 0$. then, the composition $\mu_\alpha := T \circ \alpha$ defines a fuzzy measure.*

Proof. Note first that

$$\mu_\alpha(\emptyset) := T(\alpha(\emptyset)) = T(0) = 0.$$

Then, if $A, B \in \mathcal{F}$ and $A \subseteq B$, since α is a crisp measure, $\alpha(A) \leq \alpha(B)$ but T is a monotonously increasing function, thus

$$\mu_\alpha(A) = T(\alpha(A)) \leq T(\alpha(B)) = \mu_\alpha(B).$$

Therefore, $\mu_\alpha = T \circ \alpha$ is a fuzzy measure. □

3.2.1 Types of fuzzy measures

Fuzzy measures can be classified in three groups based on the interaction between disjoint subsets. Let $\mu : \mathcal{F} \to \mathbb{R}$ be a set function and let $A, B \in \mathcal{F}$ such that $A \cap B = \emptyset$, then μ is said to be [3]:

- A **subadditive fuzzy measure**, if

  $$\mu(A \cup B) \leq \mu(A) + \mu(B).$$

  It describes a **positive interaction**, i.e., cooperation or improvement.

- A **superadditive fuzzy measure**, if

  $$\mu(A) + \mu(B) \leq \mu(A \cup B).$$

  It describes a **negative interaction**, i.e., rivalry or incompatibility.

- An **additive fuzzy measure**, if

  $$\mu(A \cup B) = \mu(A) + \mu(B).$$

  It describes a **Non- interaction**, i.e., there are not any interaction between the elements of a set and their measurements.

The most common type of monotone measure in literature is the $\lambda-$measure, also known as Sugeno $\lambda-$fuzzy measure, which can be used to compute the interdependency among the selection criteria in decision making problems [39]. Its definition is given below.

Definition 3.2.3. *Let $(X, \mathcal{F})$ be a measurable space and $\lambda \in (-1, \infty)$. A $\lambda-$**measure** is a non-negative set function $\mu_\lambda : \mathcal{F} \to [0, \infty]$ that satisfies:*

(λ1) $\mu_\lambda(X) = 1$.

(λ2) *For all $A, B \in \mathcal{F}$, if $A \cap B = \emptyset$ then*

$$\mu_\lambda(A \cup B) = \mu_\lambda(A) + \mu_\lambda(B) + \lambda\mu_\lambda(A)\mu_\lambda(B) \qquad (\lambda - \textbf{\textit{Additivity}}).$$

Note that μ_λ is a normal a fuzzy measure for each $\lambda \in (-1, \infty)$. The next proposition shows this fact and other properties.

Proposition 3.2.1. *Let μ_λ be a $\lambda-$measure on $(X, \mathcal{F})$ with $\lambda > -1$, then the following properties hold*

1. *μ_λ is a fuzzy measure.*
2. *$\mu_\lambda(A) + \mu_\lambda(A^c) = 1 - \lambda\mu_\lambda(A)\mu_\lambda(A^c)$.*
3. *$\mu_\lambda(A \cup B) = \frac{\mu_\lambda(A)+\mu_\lambda(B)-\mu_\lambda(A\cap B)+\lambda\mu_\lambda(A)\mu_\lambda(B)}{1+\lambda\mu_\lambda(A\cap B)}$.*

Proof. Let $\lambda > -1$ and let μ_λ be a $\lambda-$measure.

1. Since μ_λ is a $\lambda-$measure then $\mu_\lambda(X) = 1$. Thus,

$$\begin{aligned} 0 &= 1 - 1 = \mu_\lambda(X) - 1 = \mu_\lambda(X \cup \emptyset) - 1 \\ &= \mu_\lambda(X) + \mu_\lambda(\emptyset) + \lambda\mu_\lambda(X)\mu_\lambda(\emptyset) - 1 \\ &= 1 + \mu_\lambda(\emptyset) + \lambda\mu_\lambda(\emptyset) - 1 \\ &= \mu_\lambda(\emptyset)(1 + \lambda). \end{aligned}$$

But $\lambda > -1$; thus, $\mu_\lambda(\emptyset) = 0$.

Additionally, if $A, B \in \mathcal{F}$ such that $A \subseteq B$ then there exist $C \in \mathcal{F}$ such that $B = A \cup C$ then

$$\begin{aligned} \mu_\lambda(B) = \mu_\lambda(A \cup C) &= \mu_\lambda(A) + \mu_\lambda(C) + \lambda\mu_\lambda(A)\mu_\lambda(C) \\ &= \mu_\lambda(A) + \mu_\lambda(C)(1 + \lambda\mu_\lambda(A)) \\ &\geq \mu_\lambda(A). \end{aligned}$$

The last inequality is due to $\mu_\lambda(C)(1 + \lambda\mu_\lambda(A)) \geq 0$.
In effect, by definition $0 \leq \mu_\lambda(A) \leq 1$ then $1 - \mu_\lambda(A) \geq 0$, but $\lambda > -1$, thus $1 + \lambda\mu_\lambda(A) \geq 1 - \mu_\lambda(A) \geq 0$. Therefore, $\mu_\lambda(C)(1 + \lambda\mu_\lambda(A)) \geq 0$.

Based on this, μ_λ is monotonous and it is concluded that μ_λ is a fuzzy measure.

2. Since for all $A \in \mathcal{F}$, $A \cup A^c = X$ and $A \cap A^c = \emptyset$. Then

$$1 = \mu_\lambda(X) = \mu_\lambda(A \cup A^c) = \mu_\lambda(A) + \mu_\lambda(A^c) + \lambda\mu_\lambda(A)\mu_\lambda(A^c).$$

Hence,

$$\mu_\lambda(A) + \mu_\lambda(A^c) = 1 - \lambda\mu_\lambda(A)\mu_\lambda(A^c).$$

3. Note that, for all $A, B \in \mathcal{F}$, $A \cup B = (A \cap B^c) \cup B$ where $(A \cap B^c) \cap B = \emptyset$. Thus,

$$\mu_\lambda(A \cup B) = \mu_\lambda(A \cap B^c) + \mu_\lambda(B) + \lambda\mu_\lambda(A \cap B^c)\mu_\lambda(B). \tag{3.1}$$

Also, it is true that $A = (A \cap B) \cup (A \cap B^c)$ where this is a disjoint union for all $A, B \in \mathcal{F}$. Then,

$$\begin{aligned} \mu_\lambda(A) &= \mu_\lambda(A \cap B) + \mu_\lambda(A \cap B^c) + \lambda\mu_\lambda(A \cap B)\mu_\lambda(A \cap B^c) \\ &= \mu_\lambda(A \cap B) + \mu_\lambda(A \cap B^c)(1 + \lambda\mu_\lambda(A \cap B)). \end{aligned}$$

Thus,

$$\mu_\lambda(A \cap B^c) = \frac{\mu_\lambda(A) - \mu_\lambda(A \cap B)}{1 + \lambda\mu_\lambda(A \cap B)}. \tag{3.2}$$

Replacing (3.2) in (3.1) we conclude that

$$\begin{aligned} \mu_\lambda(A \cup B) &= \left[\frac{\mu_\lambda(A) - \mu_\lambda(A \cap B)}{1 + \lambda\mu_\lambda(A \cap B)}\right](1 + \lambda\mu_\lambda(B) + \mu_\lambda(B)) \\ &= \frac{\mu_\lambda(A) + \mu_\lambda(B) - \mu_\lambda(A \cap B) + \lambda\mu_\lambda(A)\mu_\lambda(B)}{1 + \lambda\mu_\lambda(A \cap B)}. \end{aligned}$$

□

***Remark* 19.** *Note that from the property 2 of the Proposition 3.2.1*

$$\mu_\lambda(A^c) = \frac{1 - \mu_\lambda(A)}{1 + \mu_\lambda(A)},$$

i.e., $\mu_\lambda(A^c) = N_\lambda(\mu(A))$ *where* N_λ *is the* $\lambda-$*complement.*

The following theorem gives a method to calculate the parameter λ that makes μ_λ be a fuzzy measure if the values of basic elements are given [72].

Theorem 3.2.3. *Let* X *be a finite set, i.e.,* $X = \{x_1, \ldots, x_n\}$ *where* $n \geq 2$ *and let* μ_λ *be a* $\lambda-$*measure on* $\wp(X)$. *Assume that* $1 \geq \mu_\lambda(x_i) \geq 0$ *with at least two of them being non-zero. Then the value of* λ *can be uniquely determined by*

$$\lambda + 1 = \prod_{i-1}^{n}(1 + \mu_\lambda(x_i)). \tag{3.3}$$

and

- $\lambda > 0$ *when* $\sum_{i=1}^n \mu_\lambda(x_i) < 1$.
- $\lambda = 0$ *when* $\sum_{i=1}^n \mu_\lambda(x_i) = 1$.
- $-1 < \lambda < 0$ *when* $\sum_{i=1}^n \mu_\lambda(x_i) > 1$.

Proof. To simplify the notation, let $a_i := \mu_\lambda(x_i)$ for $i = 1, \ldots, n$ and, for $k = 2, \ldots, n$, define

$$f_k(\lambda) = \prod_{i=1}^{k}(1 + a_i\lambda).$$

Note that $1 + a_i\lambda > 0$ since $a_i \in [0,1]$ and $\lambda > -1$. Furthermore, there is no loss of generality in assuming $a_1 > 0$ and $a_2 > 0$.

Also, note that $f_k(\lambda) = (1 + a_k\lambda)f_{k-1}(\lambda)$. Hence,

$$f_k'(\lambda) = a_k f_{k-1}(\lambda) + (1+a_k\lambda)f_{k-1}'(\lambda) \qquad \text{and} \qquad f_k''(\lambda) = 2a_k f_{k-1}'(\lambda) + (1+a_k\lambda)f_{k-1}''(\lambda).$$

Thus, for any $k = 2, \ldots, n$ and any $\lambda \in (-1, \infty)$, if $f_{k-1}'(\lambda) > 0$ and $f_{k-1}''(\lambda) > 0$ then so are $f_k'(\lambda)$ and $f_k''(\lambda)$. In particular, since that

$$f_2'(\lambda) = a_1(1 + a_2\lambda) + a_2(1 + a_1\lambda) > 0 \qquad \text{and} \qquad f_2''(\lambda) = 2a_1a_2 > 0,$$

the function $f_n(\lambda)$ is concave in $(-1, \infty)$. From the derivative of $f_n(\lambda)$,

$$f_n'(0) = \sum_{i=1}^{n} a_i.$$

Now, observe that $\lim_{n\to\infty} f_n(\lambda) = \infty$ but:

- If $\sum_{i=1}^{n} a_i < 1$, the curve of $f_n(\lambda)$ has a unique intersection point with the line $f(\lambda) = 1 + \lambda$ on some $\lambda > 0$
- If $\sum_{i=1}^{n} a_i = 1$, then the line $f(\lambda) = 1 + \lambda$ is just the tangent of $f_n(\lambda)$ at point $\lambda = 0$ and therefore, the curve of $f_n(\lambda)$ has no intersection point anywhere else with the line $f(\lambda) = 1 + \lambda$.
- If $\sum_{i=1}^{n} a_i > 1$, since $f_n'(\lambda) > 0$ and $f(\lambda) = 1 + \lambda \leq 0$ when $\lambda \leq -1$, the curve of $f_n(\lambda)$ must have a unique intersection point with the line $f(\lambda) = 1 + \lambda$ on some $\lambda \in (-1, 0)$.

□

The FIGURE 3.1 illustrates the reasoning done in the proof of the Theorem 3.2.3 and the example that follows illustrates its power.

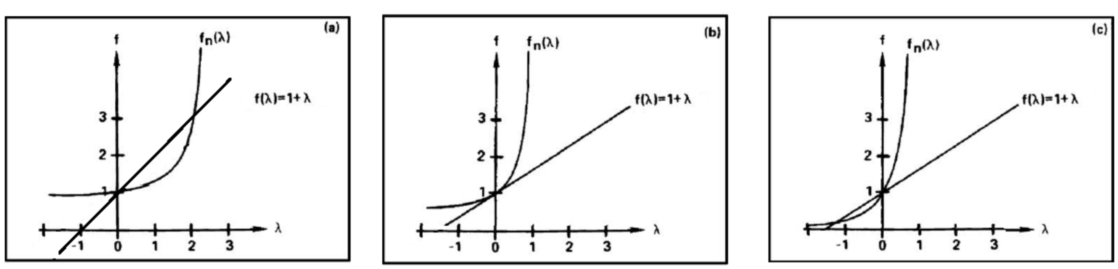

FIGURE 3.1. (modified from [72]).

Example 3.2.2. *Let* $X = \{x_1, x_2, x_3\}$ *with* $\mu_\lambda(x_1) = 0.8$*;* $\mu_\lambda(x_2) = 0.6$ *and* $\mu_\lambda(x_3) = 0.4$*. The the value of* λ *can be calculated by using equation (3.3) as follows:*

$$\begin{aligned} \lambda + 1 &= \prod_{i=1}^{3}(1 + \mu_\lambda(x_i)) \\ &= (1 + 0.8\lambda)(1 + 0.6\lambda)(1 + 0.4\lambda) \\ &= 1 + 1.8\lambda + 1.04\lambda^2 + 0.192\lambda^3. \end{aligned}$$

Thus,

$$0 = \lambda(0.8 + 1.04\lambda + 0.192\lambda^2),$$

and then conclude that

$$\lambda \in \{-4.448, -0.9283, 0\}.$$

Since $\lambda \in (-1, \infty)$*, when* $\lambda = 0$ *we obtain an additive measure.*

If $\lambda = -0.9283$*,* $\mu_\lambda(x_1) = 0.8$*;* $\mu_\lambda(x_2) = 0.6$ *and* $\mu_\lambda(x_3) = 0.4$ *then*

$$\mu_\lambda(\{x_1, x_2\}) = 0.8 + 0.6 + (-0.9283)(0.64) = 0.9554, \quad (3.4)$$

$$\mu_\lambda(\{x_1, x_3\}) = 0.8 + 0.4 + (-0.9283)(0.32) = 0.9029, \quad (3.5)$$

$$\mu_\lambda(\{x_2, x_3\}) = 0.6 + 0.4 + (-0.9283)(0.48) = 0.7772, \quad (3.6)$$

$$\mu_\lambda(X) = 1. \quad (3.7)$$

Since μ_λ *has satisfied all conditions of fuzzy measure, for* $\lambda = -0.9283$*, the* μ_λ *is a fuzzy measure.*

***Remark* 20.** *From these definitions, if* $\lambda < 0$*, the* $\lambda-$*measure* μ_λ *is subadditive, whether* $\lambda > 0$ *then* μ_λ *is superadditive and* μ_λ *is additive whenever* $\lambda = 0$*. The* FIGURE *3.2 shows the relationship between* $\lambda-$*measures and the additive, subadditive and superadditive measures.*

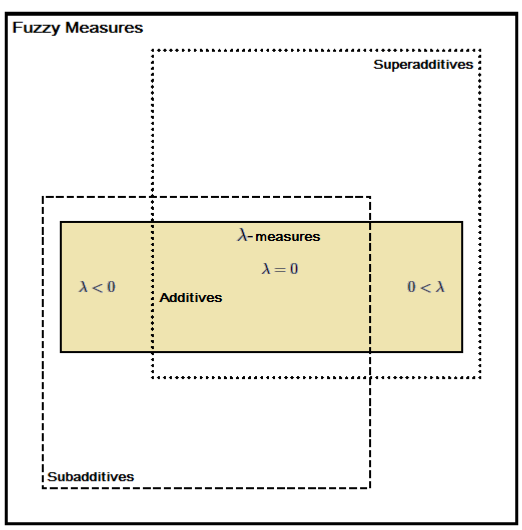

FIGURE 3.2. (modified from [3]).

The following are examples of fuzzy measures which have become an important basis for some mathematical theories like: probability theory, evidence theory and necessity theory.

3.2.1.1 Probability Measure (Additive Fuzzy Measure)

A probability is just a real function whose domain is a set of sets and satisfies some properties of a crisp measure itself.

Definition 3.2.4. *Let $(X, \mathcal{F})$ be a measurable space. A* ***probability measure*** *is a function $P : \mathcal{F} \to \mathbb{R}$ satisfying:*

1. *$P(A) \geq 0$, for all $A \in \mathcal{F}$.*
2. *$P(X) = 1$.*
3. *Whether $\{E_n\}_{n \in \mathbb{Z}^+} \subseteq \mathcal{F}$ is a disjoint family, then*

$$P\left(\bigcup_{n \in \mathbb{Z}^+} E_n\right) = \sum_{n=1}^{\infty} P(E_n).(\sigma - \textbf{\textit{additivity}}).$$

We denote the probability space by $(X, \mathcal{F}, P)$.

We affirm that every probability measure P is an additive fuzzy measure.

Proposition 3.2.2. *Let P be a probability measure then P is a fuzzy measure; indeed, P is an additive fuzzy measure.*

Proof. We must show that $P(\emptyset) = 0$ and $P(A) \leq P(B)$ whenever $A \subseteq B$.
Note that $X = X \cup \emptyset$, then

$$1 = P(X) = P(X \cup \emptyset) = P(X) + P(\emptyset) = 1 + P(\emptyset).$$

Hence, $P(\emptyset) = 0$.

On the other hand, if $A \subseteq B$ then $B = A \cup (B \setminus A)$, then $P(B) = P(A) + P(B \setminus A)$ but, $P(B \setminus A) \geq 0$, hence $P(B) \geq P(A)$.

Therefore, P is an additive fuzzy measure since it satisfies the $\sigma-$additivity. □

The following two examples of fuzzy measures are the basis of the theory of evidence, which was developed by Dempster [17] and extended by Shafer [57] in order to represent ignorance. This theory focuses in the believability assigned to an event by keeping the experience and point of view of the decision maker in mind; in contrast with probability theory which assumes the existence of values associated to some events, whether or not the observer knows the real value of the probability [4].

3.2.1.2 Plausibility Measure (Subadditive Measure)

Plausibility measure is an important example of subadditive measure which is used to measure grades of verosimilitude (i.e., what seems to be true and believable) since they represent the maximum belief in an hipotesis as a result of an evidence [3].

Definition 3.2.5. *Let $(X, \mathcal{F})$ be a measure space. A* ***plausibility measure*** *is a function $\mu_P : \mathcal{F} \to [0, 1]$ satisfying*

(Pl 1) $\mu_P(\emptyset) = 0$.

(Pl 2) $\mu_P(X) = 1$.

(Pl 3) *If* $\{A_i\}_{i=1}^n \subseteq \mathcal{F}$ *then*

$$\mu_P\left(\bigcap_{1\leq i\leq n} A_i\right) \leq \sum_{1\leq i\leq n} \mu_P(A_i) - \sum_{j\leq k} \mu_P(A_j \cup A_k) + \cdots + (-1)^{n+1}\mu_P\left(\bigcup_{1\leq i\leq n} A_i\right).$$

We affirm that every plausibility measure is a fuzzy measure but the proof of this sentence will be given further. Another thing to keep in mind is that when $n = 2$, the property **(Pl3)** can be written as

$$\mu_P\left(A_1 \cap A_2\right) \leq \mu_P(A_1) + \mu_P(A_2) - \mu_P(A_1 \cup A_2).$$

Hence, whether $A_1 \cap A_2 = \emptyset$,

$$\mu_P(A_1 \cup A_2) \leq \mu_P(A_1) + \mu_P(A_2).$$

Thus, μ_P is an subadditive fuzzy measure such that for all $A \in \mathcal{F}$,

$$1 = \mu_P(X) = \mu_P(A \cup A^c) \leq \mu_P(A) + \mu_P(A^c).$$

3.2.1.3 Believability Measure (Superadditive Measure)

The believability measure is an example of superadditive measures which is important in the evidence theory. This theory assigns, from the point of view and experience of the decision-maker, believability (credibility) to an event that occurs or that had occurred [3].

Definition 3.2.6. *Let* $(X, \mathcal{F})$ *be a measure space. A function* $\mu_B : \mathcal{F} \to [0,1]$ *is a* ***believability measure*** *when it accomplishes*

(Bel 1) $\mu_B(\emptyset) = 0$.

(Bel 2) $\mu_B(X) = 1$.

(Bel 3)

$$\mu_B\left(\bigcup_{1\leq i\leq n} A_i\right) \geq \sum_{1\leq j\leq n} \mu_B(A_j) - \sum_{j\leq k} \mu_B(A_j \cap A_k) + \cdots + (-1)^{n+1}\mu_B\left(\bigcap_{1\leq i\leq n} A_i\right).$$

The image of an element $A \in \mathcal{F}$, $\mu_B(A)$, is called **belief degree in** A and represents the minimum belief in the hipotesis A as a result of a test and, $\mu_B(A^c)$ is termed **doubt degree in** A and represents the minimum belief in the negation of hipotesis A as a result of a test [3]. When $n = 2$, the property **(B3)** can be written as

$$\mu_B\left(A_1 \cup A_2\right) \geq \mu_B(A_1) + \mu_B(A_2) - \mu_B(A_1 \cap A_2).$$

Hence, whether $A_1 \cap A_2 = \emptyset$,

$$\mu_B(A_1 \cup A_2) \geq \mu_B(A_1) + \mu_B(A_2).$$

Thus, μ_B is a superadditive fuzzy measure and for each $A \in \mathcal{F}$

$$1 \geq \mu_B(A) + \mu_B(A^c).$$

Remark* 21.** *Another way to define a believability measure is as follows. First, given a non-empty set X, let p be a discrete probability measure on $(\wp(X), \wp(\wp(X)))$ such that $p(\{\emptyset\}) = 0$ and consider a* ***basic probability assignment *m defined as the function*

$$\begin{aligned} m : \wp(X) &\to [0,1] \\ E &\mapsto p(\{E\}). \end{aligned}$$

If this function satisfies that $m(\emptyset) = 0$ and $\sum_{E \in \wp(X)} = 1$ then, for all $E \in \wp(X)$, the formula

$$\mu_B(E) := \sum_{F \subset E} m(F)$$

determines a ***belief measure****. Also, for $E \in \wp(X)$, a* ***plausibility measure*** *can be determined by the formula*

$$\mu_P := \sum_{F \cap E \neq \emptyset} m(F).$$

Now, it will be shown that a believability measure μ_B is a fuzzy measure.

Proposition 3.2.3. *Every believability measure μ_B is a fuzzy measure.*

Proof. By definition $\mu_B(\emptyset) = 0$. To prove the monotonicity, let $A, B \in \mathcal{F}$ such that $A \subseteq B$ and let $C := B \setminus A$, then

$$\mu_B(B) = \mu_B(A \cup C) \geq \mu_B(A) + \mu_B(C) \geq \mu_B(A).$$

□

For each fuzzy measure μ we can define its dual μ_d through the following relation

$$\mu_d(A) := \mu(X) - \mu(A^c), \text{ for all } A \in \mathcal{F}.$$

In particular, we have the following equality between plausibility measures μ_P and believability measures.

Proposition 3.2.4. *Let μ_P be a plausibility measure and let μ_B be a believability measure on $(X, \mathcal{F})$, then:*

1. *For all $A \in \mathcal{F}$,*
$$\mu_P(A) = 1 - \mu_B(A^c).$$
2. *or all $A \in \mathcal{F}$,*
$$\mu_P(A) \geq \mu_B(A),$$

Proof. Let μ_P be a plausibility measure on $(X, \mathcal{F})$, μ_B be a believability measure on $(X, \mathcal{F})$ and $A \in \mathcal{F}$, then:

1. From the alternative definition of believability measure and plausibility measure,

$$\begin{aligned}\mu_B(A) &= \sum_{F\subset A} m(F)\\ &= \sum_{F\subset X} m(F) - \sum_{F\not\subseteq A} m(F)\\ &= 1 - \sum_{F\cap A^c\neq\emptyset} m(F)\\ &= 1-\mu_P(A^c).\end{aligned}$$

2. Since

$$\mu_P(A)+\mu_P(A^c)\geq 1 \qquad \text{and} \qquad 1-\mu_P(A^c)=\mu_B(A)$$

then

$$\mu_P(A)\geq 1-\mu_P(A^c)=\mu_B(A).$$

Hence,

$$\mu_P(A)\geq\mu_B(A),\ \text{for all } A\in\mathcal{F}.$$

□

As a consequence of the Proposition 3.2.4, the following result arises.

Corolary 3.2.4. *Every plausibility measure μ_P is a fuzzy measure.*

Proof. Let μ_P be a plausibility measure. By definition, $\mu_P(\emptyset)=0$. Also if $A\in\mathcal{F}$ then $0<1=\mu_P(X)=\mu_P(X\cup A)\leq\mu_P(X)+\mu_P(A)$; hence, $\mu_P(A)\geq 0$. Now, if $A,B\in\mathcal{F}$ and $A\subseteq B$ then:

$$\mu_P(A)=1-\mu_B(A^c)\leq 1-\mu_B(A^c)=\mu_P(B).$$

Hence, μ_P is a fuzzy measure. □

To finish this section, an example for these last measures is given.

Example 3.2.3 (Believability and Plausibility measure). *Let $X=\{a,b\}$, $\mathcal{F}=\wp(X)$ and $\mu:\wp(X)\to[0,\infty)$ be a function defined by*

$$\mu(A)=\begin{cases}0 & A=\emptyset\\ t & A\in\{\{a\},\{b\}\}\\ 1 & A=X\end{cases}$$

where $t\in(0,1)$ fixed. Then, μ is monotonous and a fuzzy measure too. Note that μ is a believability measure whenever $t\in(0,\frac{1}{2})$ and it is a plausibility measure when $t\in(\frac{1}{2},1)$. If $t=\frac{1}{2}$, then we obtain a probability measure [3].

3.2.2 Transformations for fuzzy measures

In order to define two useful transformations, we are going to deal with the vector space of set functions. First we will say that a **transformation** is a mapping $\Psi:\mathbb{R}^{2^N}\to\mathbb{R}^{2^N}$ that assigns to a set function μ another set function Ψ^{μ}. If the transformation is linear and invertible then it induces a basis of the vector space of the set functions; conversely, each basis induces a linear invertible transformation [28].

The first transformation of interest is known as the **Interaction transform** which enables the interpretation of fuzzy measures in multicriteria decision making context [28]. The interaction transform of μ can be expressed through its Möbius transform as follows:

$$I^{\mu}(A) := \sum_{B \supseteq A} \frac{1}{b-a+1} m^{\mu}(B) \qquad A \in \wp(X).$$

where a and b are cardinalities of subsets A and B, respectively and,

$$m^{\mu}(B) := \sum_{D \supseteq B} (-1)^{|B \setminus D|} \mu(D), \qquad B \in \wp(X).$$

The second transform of interest is the **Fourier transform** which is known in computer science and defined by

$$F^{\mu}(A) = \frac{1}{2^n} \sum_{K \subseteq X} (-1)^{|A \cap K|} \mu(K);$$

where, $|X| = n$, $A \in \wp(X)$. The elements of the basis formed by this transformation are called **parity functions** and form an orthogonal basis of the set of square integrable functions on $[0,1]$ [28], where the **inner product** is defined by

$$\langle \mu, \xi \rangle := \frac{1}{2^n} \sum_{A \subseteq X} \mu(A)\xi(A);$$

where $\mu(A) = \sum_{B \subseteq A} m^{\mu}(B)$ and $\mu(A) = \sum_{B \subseteq A} m^{\xi}(B)$. The **convolution product** turns into an ordinary product, i.e.,

$$F^{\mu \star \xi} = F^{\mu} F^{\xi};$$

where the convolution of two set functions is defined by

$$\mu \star \xi := \frac{1}{2^n} \sum_{T \subseteq X} \mu\left((A \cup T) \setminus (A \cap T)\right) \xi(T).$$

A fuzzy measure μ is $k-$**additive** $(1 \leq k \leq n)$ if its interaction transform I^{μ} vanishes for subsets of more than k elements and there is $B \subseteq X$ such that $|B| = k$ and $I^{\mu}(B) \neq 0$. The interaction transform has a clear interpretation in the context of multicriteria decision making, therefore $k-$additive fuzzy measures are of particular interest. Indeed, $k-$additive fuzzy measures are families of fuzzy measures which are of polynomial complexity instead of the exponential complexity of general fuzzy measures [28].

3.2.3 Another issues on Fuzzy Measures

In the context of decision making, by assuming we have a collection of criteria of interest, we can use aggregation functions to combine several numerical values into a single one. Common aggregator operators like the Arithmetic mean, the Weighted mean or the mode are not capable to find interaction between criteria but, through t-norms and t-conorms like max or min we can express the structural relationship between criteria [3, 27, 39]. Also, we can obtain a very general formulation with the aid of fuzzy measures to represent these relationships. The following theorem supports this last sentence.

Theorem 3.2.5 (**Fundamental Theorem on Aggregation Measures** [76]). *Assume that* $\mu_1, \mu_2, \dots, \mu_n$ *are fuzzy measures on the measurable space* $(X, \mathcal{F})$*. If* $Agg : [0,1]^n \to [0,1]$ *is an aggregation function, that is,* Agg *satisfies:*

1. $Agg(0, \dots, 0) = 0$.
2. $Agg(1, \dots, 1) = 1$.
3. $Agg(a_1, \dots, a_n) \geq Agg(b_1, \dots, b_n)$ *whenever* $a_i \geq b_i$ *for all* $i = 1, 2, \dots, n$.

Then, the set function $\mu : \wp(X) \to \mathbb{R}$ *defined by*

$$\mu(A) := Agg(\mu_1(A), \mu_2(A), \dots, \mu_n(A)). \text{ for all } A \in \wp(X)$$

is itself a fuzzy measure.

3.2.3.1 The Fuzzy Integral

Some of the applications of fuzzy measures and fuzzy integrals lay on branches of engineering or social sciences like: decision making, information recovery, data modeling, analysis of patterns and attitudesand, classification [3]. Fuzzy integrals are integrals of a real function with respect to a fuzzy measure, by analogy with Lebesgue integral which is defined with respect to an additive measure. There are several definitions of fuzzy integrals, among which the most representative are the Sugeno's integral [64] and Choquet's integral [13]. The former is based on linear operators and has been applied in the subjective evaluation of phenomena meanwhile the latter is based on no linear operators and has been used to represent statistical measures like mean, median and $L-$estimations [3].

Let $(X, \mathcal{F})$ be a measurable space and $\mu : \mathcal{F} \to [0, \infty)$ a continuous fuzzy measure then we say that $\mathcal{G}$ represents **the class of all finite non-negative measurable functions defined on** $(X, \mathcal{F})$ and that $f : (X, \mathcal{F}_X) \to (Y, \mathcal{F}_Y)$ is a measurable function whether $f^{-1}(B) \in \mathcal{F}_X$ for all $B \in \mathcal{F}_Y$, i.e., $f^{-1}(\mathcal{F}_Y) \subseteq \mathcal{F}_X$ [5]. Additionally, for each $f \in \mathcal{G}$ we define, for $\alpha \in (0, \infty)$:

$$f_\alpha = \{x \in X \mid f(x) \geq \alpha\}, \tag{3.8}$$
$$f_{\alpha^+} = \{x \in X \mid f(x) > \alpha\} \tag{3.9}$$

Since the rank of the considered functions is $[0, \infty)$, it is convenient the following notation used in Arenas' paper [3]:

$$\inf_{x \in \emptyset} f(x) = \infty.$$

Definition 3.2.7. *Let* $A \in \mathcal{F}$ *and* $f \in \mathcal{G}$*. The* ***Sugeno's integral of*** f ***on*** A ***with respect to*** μ *is defined by*

$$\int_A f d\mu := \sup_{\alpha \in [0,\infty)} \min\{\alpha, \mu(A \cap f_\alpha)\}. \tag{3.10}$$

Fuzzy measures are considered to be a generalization of classical measures. The Sugeno's integral has been used to obtain the expected value of a random variable with respect to a probability measure meanwhile the Choquet's integral was proposed in 1954 as a functional with respect to a fuzzy measure in order to study some problems in mechanics [15, 28]. The Choquet's integral arises in natural way from Lebesgue's integral and has been successfully applied in practical problems related to classification, image processing, decision making under uncertainty and data modeling [3].

Definition 3.2.8. *Let $(X, \mathcal{F})$ be a measurable space and $f \in \mathcal{G}$, the collection of all finite non-negative measurable functions defined on $(X, \mathcal{F})$. The* ***Choquet's integral of f with respect to the fuzzy measure μ in a measurable set*** *A is defined by*

$$\int_A f d\mu = \int_0^\infty \mu(A \cap f_\alpha) d\alpha. \tag{3.11}$$

In order to show that if μ is the Lebesgue measure, we must make that Choquet integral coincide with the Lebesgue integral and we need the following alternative definition of the Choquet integral.

Definition 3.2.9 (Alternative definition of the Choquet integral). *Let $0 = a_0 \leq a_1 \leq a_2 \leq \cdots \leq a_n$ and $A_i := \{x \mid f(x) \geq a_i\}$, for $i = 1, 2, \ldots, n$. We have that $A_n \subseteq A_{n-1} \subseteq \cdots \subseteq A_1$. We define the Choquet integral of a function of the form*

$$s = \sum_{i=1}^{n} (a_i - a_{i-1}) \chi_{A_i}$$

as

$$\int s d\mu = \sup_{\substack{s \in S \\ s \leq f}} \int s d\mu. \tag{3.12}$$

The following proposition says that both definitions given for Choquet integral are equivalent.

Proposition 3.2.5. *Let μ be a fuzzy measure and f be a measurable function on $(X, \mathcal{F})$, then:*

$$\int_0^\infty \mu(\{x \mid f(x) \geq \alpha\}) d\alpha = \sup_{\substack{s \in S \\ s \leq f}} \int s d\mu.$$

Proof.

$$\begin{aligned} \sup_{\substack{s \in S \\ s \leq f}} \int s d\mu &= \sup_{s \leq f} \sum_{i=1}^{n} \mu(A_i)(a_i - a_{i-1}) && (3.13) \\ &= \int_0^\infty \mu(\{x \mid f(x) \geq \alpha\}) d\alpha && (3.14) \end{aligned}$$

The latter equality is because of being a Riemann integral. □

Now, the result mentioned above can be proved.

Theorem 3.2.6. *If μ is the Lebesgue measure then the Choquet's integral coincides with a Lebesgue integral.*

Proof. If μ is the Lebesgue measure then, from the alternative definition 3.12 we have that

$$\mu(A_i) = \mu(A_{i+1}) + \mu(A_i \setminus A_{i+1}).$$

Then,

$$\int s d\mu = \sum_{i=1}^{n} (a_i - a_{i-1})\mu(A_i) \quad (3.15)$$

$$= \sum_{i=1}^{n} a_i(\mu(A_i - \mu(A_{i+1}) \quad (3.16)$$

$$= \sum_{i=1}^{n} a_i(\mu(A_i \setminus A_{i+1}) \quad (3.17)$$

which converges to the Lebesgue integral and we get

$$\int s d\mu = \int f d\mu.$$

□

***Remark* 22.** *The Choquet's integral is similar to the Sugeno's one because, in spite of μ is not necessarily additive, it is generally nonlinear with respect to its "integrand".*

There are other kinds of integral defined with respect to fuzzy measures. The **Pseudo-additive integrals** and **fuzzy $t-$conorm integrals** which basically use as basis operators a continous $t-$conorm S and a uninorm U which is distributive with respect to S in the following sense [28]:

$$U(x, S(y, z)) = S(U(x, y), U(x, z));$$

for all $x, y, z \in [0, 1]$ such that $S(y, z) < 1$.

The **universal integral** is a functional $\mathcal{I} : M(X, \mathcal{F}) \times \mathcal{A}(X, \mathcal{F}) \to [0, \infty]$, where $M(X, \mathcal{F})$ is the set of all fuzzy measures on $(X, \mathcal{F})$ measurable space and, $\mathcal{A}(X, \mathcal{F})$ is the collection of $\mathcal{F}-$measurable applications (functions), satisfying the following axioms:

1. For any measurable space $(X, \mathcal{F})$, its restriction to $D := M(X, \mathcal{F}) \times \mathcal{A}(X, \mathcal{F})$ is nondecreasing in each place.
2. There exists a pseudo-multiplication $\otimes$ such that for all $(\mu, c\Xi_A) \in D$, $\mathcal{I}(\mu, c\Xi_A) = c \otimes \mu(A)$.
3. If $G_{\mu,f} = G_{\mu',f'}$ then $\mathcal{I}(\mu, f) = \mathcal{I}(\mu', f')$.

Sugeno's and Choquet's integrals are universal integrals [28].

3.3 Possibility Measure

In the stochastic theory, the available information is treated by means of probability density functions; whereas in FuSeTh, this information might be modeled through membership functions which can be built by experts who indicate which elements have the highest or the lowest weight according to each expert's knowledge or judgement [15]. In FuSeTh, a concept which has been widely studied is the possibility [3, 15] since Zadeh introduced it but, due to its comparisons with probability theory, it has been frequently questioned to this day. As we mentioned before, Shafer in 1976 introduced believability and plausibility measures to support its evidence theory where one branch of it is the possibility theory which was studied by Zadeh [93]. In the possibility theory possibility measure and necessity one are defined as follows:

$$\pi(A \cup B) = \max\{\pi(A), \pi(B)\} \qquad (\textbf{possibility}).$$

$$\eta(A \cup B) = 1 - \Pi(A^c) \qquad (\textbf{necessity}).$$

Also, it is remarked that $\eta(A \cap B) = \min\{\eta(A), \eta(B)\}$; in other words, a possibility measure is a plausibility measure with the condition $\eta(A \cup B) = \max\{\eta(A), \eta(B)\}$ and a necessity measure is a believability measure with the additional condition $\pi(A \cap B) = \min\{\pi(A), \pi(B)\}$.

Fuzzy Measure Theory is the mathematical theory for possibility distributions and a possibility distribution is a normal fuzzy subset of a crisp set X which is adopted as the universe of discourse [6]. One of the discussions is related to the naive sentence "such a fact is possible but unlikely" which suggest that a notion for possibility measure must satisfy the principle of consistence [15]: *"If $(X, \mathcal{F})$ is a measurable space, $A \in \mathcal{F}$, π is a possibility measure and P is a probability measure; then: $\pi(A) \geq P(A)$".*

Definition 3.3.1 (Possibility distribution). *A possibility distribution π of a fuzzy variable x on a set of possible situations X is a fuzzy set $\pi : X \to [0, 1]$ such that exist an element $x_0 \in X$ with $\pi(x_0) = 1$.*

A possibility distribution is a function that represents the knowledge of an expert in a flexible way and aims to distinguish between what is expected and what is surprising, what is the normal course of things from what is not. Thus, $\pi(u)$ means the degree of possibility of the assignment $x = u$ which can be rejected as impossible when $\pi(u) = 0$ or can be totally possible whenever $\pi(u) = 1$.

Let us consider a crisp set $A \subseteq X$ then we can ask about what is the possibility degree of $x \in A$. Indeed, a natural question that arises is: Does x necessarily belongs to A? Well, the following definition aims to help us to answer it.

Definition 3.3.2. *The possibility degree of x to belong to A is*

$$\pi(A) := \sup_{x \in A} \pi(x),$$

and, the necessity degree is

$$\eta(A) := \inf_{x \in A} N(\pi(x)) = \inf_{x \in A} 1 - \pi(A).$$

Note that in this case $N(\cdot)$ is the Standard negation.

As we said in the previous section, possibility theory involves two concepts: possibility measure and necessity measure. They are defined below.

Definition 3.3.3. *Let $(X, \mathcal{F})$ be a measurable space. A **possibility measure** is a function $\pi : \mathcal{F} \to [0, 1]$ satisfying:*

(Poss 1) $\pi(\emptyset) = 0$.

(Poss 2) $\pi(X) = 1$.

(Poss 3) $A \subseteq B$ *implies* $\pi(A) \leq \pi(B)$.

(Poss 4) *Let* $I = \{1, 2, \ldots, n\}$ *be a set of indexes, then*

$$\pi\left(\bigcup_{i \in I} A_i\right) = \sup_{i \in I} \pi(A_i) = \max_{1 \leq i \leq n} \pi(A_i).$$

The previous definition and the next one were given by Garmendia in his work, we suggest to read [26] for more details.

Definition 3.3.4. *Let* $(X, \mathcal{F})$ *be a measurable space. A* ***necessity measure*** *is a function* $\eta : \mathcal{F} \to [0, 1]$ *satisfying:*

(Nec 1) $\eta(\emptyset) = 0$.

(Nec 2) $\eta(X) = 1$.

(Nec 3) $A \subseteq B$ *implies* $\eta(A) \leq \eta(B)$.

(Nec 4) *Let* I *be a set of indexes, then*

$$\eta\left(\bigcap_{i \in I} A_i\right) = \inf_{i \in I} \eta(A_i) = \min_{1 \leq i \leq n} \eta(A_i).$$

Possibility measures and necessity measures are fuzzy measures, the former satisfies the subadditivity property and for the latter it holds the superadditivity property.

Proposition 3.3.1. *Let* $(X, \mathcal{F})$ *be a measurable space and* $\pi : \mathcal{F} \to [0, 1]$ *and* $\eta : \mathcal{F} \to [0, 1]$ *be a possibility and necessity measure, respectively, then*

1. $\eta(A) + \eta(A^c) \leq 1 \leq \pi(A) + \pi(A^c)$.
2. $\eta(A) = 0$ *implies* $\pi(A) = 1$.
3. $\eta(A) = 0$ *whenever* $\pi(A) \geq 1$.

Proof. 1. Since $0 = \min\{\eta(A), \eta(A^c)\}$ and $\max\{\eta(A), \eta(A^c)\} \leq \eta(X)$, then

$$\min\{\eta(A), \eta(A^c)\} + \max\{\eta(A), \eta(A^c)\} \leq \eta X = 1.$$

Hence,

$$\eta(A) + \eta(A^c) \leq 1.$$

On the other hand,

$$1 = \max\{\pi(A), \pi(A^c)\} \leq \max\{\pi(A), \pi(A^c)\} + \min\{\pi(A), \pi(A^c)\} = \pi(A) + \pi(A^c).$$

Thus, we have proved the desired.

2. Since, $0 = \min\{\eta(A), \eta(A^c)\}$ and $\eta(A) \geq 0$ we can conclude that $\eta(A^c) = 0$ but, $\eta(A^c) = 1 - \pi(A)$. Hence, $\pi(A) = 1$.

3. Since, $1 = \max\{\pi(A), \pi(A^c)\}$ and $\pi(A) \geq 1$ we can conclude that $\pi(A^c) = 1$ but, $\pi(A^c) = 1 - \eta(A)$. Hence, $\eta(A) = 0$.

□

Remark 23. *Whether it is given a possibility measure π then, it introduces a possibility distribution function Φ_π on X through its constraint to the elements of X, i.e.,*

$$\Phi_\pi(x) = \pi(\{x\}).$$

In fact, since $\pi(X) = 1$ and since we can write $X = \bigcup_{x \in X} \{x\}$, we obtain $\pi(X) = \sup_{x \in X} \pi(\{x\}) = 1$. So there exist $x_0 \in X$ with $\pi(\{x_0\}) = 1$ or equivalently, $\pi(x_0) = 1$

Remark 24. *Note that if $A \in \mathcal{F}$ then:*

$$1 = \pi(X) = \pi(A \cup A^c) = \max\{\pi(A), \pi(A^c)\}.$$

$$0 = \eta(\emptyset) = \pi(A \cap A^c) = \min\{\eta(A), \eta(A^c)\}.$$

Another thing to remark is that a necessity measure can be obtained from a possiblity measure by means the following equation [93]:

$$\eta(A) = 1 - \pi(A^c).$$

$$\pi(A^c) = N(\eta(A)).$$

Proposition 3.3.2. *Let $\Phi : X \to [0, 1]$ be a possibility distribution then, it introduces a possibility measure on X as follows:*
For all $A \in \mathcal{F}$,

$$\pi(A) := \begin{cases} \sup_{x \in A} \Phi(x) & A \neq \emptyset \\ 0 & A = \emptyset. \end{cases}$$

Proof. For the proof, we adopt the convention $\pi(\emptyset) = 0$ and because π is a distribution of possibility then $\pi(X) = \sup_{x \in X} \pi(\{x\}) = 1$ [15].

Let $\{A_i\}_{i \in I} \subseteq \mathcal{F}$ then

$$\begin{aligned} \pi\left(\bigcup_{i \in I} A_i\right) &= \sup\{\pi(x) \mid x \in \bigcup_{i \in I} A_i\} \\ &= \sup_{i \in I} \sup_{x \in A_i} \pi(x) \\ &= \sup_{i \in I} \pi(A_i). \end{aligned}$$

□

In many practical problems is essential the existence of a way of transform probabilities in possibilities and vice versa. All adopted methods to do it have in common that they satisfy the consistency principle, i.e., for all $A \in \mathcal{F}$, $P(A) \leq \pi(A)$.

The simplest transformation given in literature are [15]:

$$\pi(x_i) = \frac{P(x_i)}{P(x_1)}$$

and

$$P(x_i) = \frac{\pi(\omega_i)}{\sum_{j=1}^{n} \pi(\omega_j)},$$

where $X = \{x_1, x_2, \ldots, x_n\}$ such that

$$1 = \pi(x_1) \geq \pi(x_2) \geq \cdots \geq \pi(x_n)$$

and

$$1 = P(x_1) \geq P(x_2) \geq \cdots \geq P(x_n).$$

3.3.1 Possibility Theory vs Probability Theory

Uncertainty-based information has been represented and dealt with terms of classical set theory or probability theory; however there are some limitations and disadvantages with the probabilistic method but now it can be well understood in terms of FuSeTh and tools of Evidence theory (also know as Dempster-Shafer theory) like possibility and necessity measures.

Fuzzy measure theory is interesting due to its three special branches: Probability theory, Evidence theory and Possibility theory. Probability theory is a tool for formalizing uncertainty in situations where frequencies are known or where evidence is based on results of series of independent random experiments, provides a method for representing incomplete information but the probabilistic method has some disadvantages, for instance, the probabilistic methods always require prior probabilities which are very hard to find out a priori [33]. Evidence is often associated with only one possible event. Possibility theory is a mathematical theory for dealing with uncertainty associated with the handling of incomplete information and Zadeh introduced it to give a formal reasoning on imprecise knowledge and make possible to deal with uncertainties related to. Possibility theory is comparable to probability one since both are based on set functions. If we have enough information about uncertainties and accurate predictive models, then probability is advantageous. When making decisions under limited information it may be useful to consider both probability and possibility of failure of the system. This section compiles some of the main ideas given by Nayal [1] whose document discusses several differences and similarities between probability measures and possibility.

Among the similarities we find that both probability and possibility measures are defined on a set $\mathcal{F} \subseteq \wp(X)$ and they are used to model uncertainty. On the other hand, their differences are much more visible; the Probability is a tool for formalizing incomplete information since it is a measure of the frequency of occurrence of an event where the value of each probability distribution are required to add to 1, it makes probability measure be an additive one meanwhile the possibility is a tool for formalizing uncertainty due to it is a measure used to quantify the meaning of an event whose its distribution requires that the largest values be equal to 1, it makes possibility measures be subadditive ones.

The table 3.1 presents some of the more representative mathematical properties of probability and possibility theory.

Probability Theory	**Possibility Theory**
One type of measure as basis: Probability measure (P).	Two types of measures as basis: Possibility measure (π) and Necessity measure (η).
It is held additivity, i.e., $P(A \cup B) = P(A) + P(B) - P(A \cap B)$,	$\pi(A \cup B) = \max\{\pi(A), \pi(B)\}$. $\eta(A \cup B) = \min\{\eta(A), \eta(B)\}$.
Unique representation of P by a probability distribution function $P : X \to [0, 1]$ via the formula $P(A) = \sum_{x \in A}$.	Unique representation of π by a possibility distribution function $\pi : X \to [0, 1]$ via the formula $\pi(A) = \max_{x \in A} \pi(x)$.
It is normalized by $\sum_{x \in X} p(x) = 1$.	It is normalized by $\max_{x \in X} \pi(x) = 1$.
Total ignorance: For all $x \in X$ $p(x) = \frac{1}{\lvert X \rvert}$. $P(A) + P(A^c) = 1$	Total ignorance: For all $x \in X$ $\pi(x) = 1$. $\pi(A) + \pi(A^c) \geq 1$, $\eta(A) + \eta(A^c) \leq 1$, $\max\{\pi(A), \pi(A^c)\} = 1$, $\min\{\eta(A), \eta(A^c)\} = 0$.

TABLE 3.1. Basic mathematical properties of the probability theory and the possibility theory.

Chapter 4

Theory of Fuzzy Computation

Numbers have played an important role in our culture and language but there are situations where we might have to do complex computations that we cannot perform in our heads. It motivates the invention of computing devices like the Turing machine, which is not only considered the archetypal conceptual computing device but it has become in the basis of modern computability theory [65]. Here we pretend to present these type of devices in terms of the language of FuSeTh due to many researchers tend to confuse the notion of fuzzy computation with fuzzy expert systems or other applications of the FuSeTh; see [65] for more details about.

The main purpose of a given field of study is to contribute to knowledge [58] but a new theory should not ignore the predictions or the results delivered by older theories [66]. Fuzzy sets were introduced to model cases or objects from another perspective and so, a theory of fuzzy computation (TFC) is a theory whose purpose is to allow us to compute in environments where one cannot precisely measure and it should enrich current theories by including vagueness into them [65], in this part of the document we intend to seed the idea that fuzzy computation should be considered an interesting field of study.

Fuzziness is not a linguistic phenomenon but a mathematical model of vagueness and the fusion of computability theory with FuSeTh demands a good knowledge of both theories [65]. In Chapter 1, the main concepts and results of FuSeTh are presented but here we deal only with a model of computation where vagueness is described using FuSeTh: The Fuzzy Turing Machine.

4.1 The Classical Turing Machine

In this section, we present some definitions and results that we reckon that help us to understand the theory related to Turing machines.

We start by saying that an **alphabet** Σ is a nonempty finite set whose elements are called **symbols**. Over the alphabet Σ, any finite sequence of symbols will be called **string** (or **word**), the empty word will be denoted by λ and the set of all strings over the alphabet Σ (including the empty string) will be denoted by Σ^*. The concatenation of $u, v \in \Sigma^*$, denoted uv, is defined as follows:

$$uv = \begin{cases} u & v = \lambda, \\ u_1u_2\cdots u_m v_1 v_2 \cdots v_n & u = u_1u_2\cdots u_m,\ v = v_1v_2\cdots v_n. \end{cases}$$

From this, for each $n \in \mathbb{N}$ we can define the powers of a string recursively as:

$$u^n = \begin{cases} \lambda & n = 0, \\ \underbrace{uu\cdots u}_{n \text{ times}} & n \geq 1. \end{cases}$$

Also, over an alphabet Σ, any finite or infinite subset L of Σ^* (i.e., $L \subseteq \Sigma^*$) is termed **language**. Some operators over languages are based on set operations such that the union, the intersection, the difference and the complement; others like concatenation, power and the Kleene star are extensions of the string operations to languages and they are termed **linguistic operations**.

Given an alphabet Σ, we can find the **regular languages** which are languages generated from the basic languages $\emptyset$, $\{\lambda\}$, $\{a\}$, $a \in \Sigma$ and the regular operations: union, concatenation and Kleene star. It is important to observe that every finite language $L = \{w_1, w_2, \ldots, w_n\}$ is regular. In order to simplify the description of the regular languages, we use **regular expressions** which may not be unique. They are defined as follows:

1. Basic regular expressions:
   $\emptyset$ is a regular expression that represents the language $\emptyset$.
   λ is a regular expression that represents the language $\{\lambda\}$.
   a is a regular expression that represents the language $\{a\},\ a \in \Sigma$.
2. Whether R and S are regular expressions over Σ, then their union $(R \cup S)$, their concatenation (RS or SR) and their closure (R^* or S^*) are regular expressions too.

The following examples let to understand some of the previous definitions.

Example 4.1.1. *When it is considered the alphabet $\Sigma = \{0, 1\}$, the language whose elements are all the even length strings and alternating zeros and ones is $L = (01)^+ \cup (10)^+$.*

Example 4.1.2. *If we consider the alphabet $\Sigma = \{0, 1, 2\}$, the language whose elements are all the strings not having two ones consecutive is $L = ((0 \cup 2)^* 1 (0 \cup 2)^+ 1 (0 \cup 2)^*)^*$.*

The following result can be used to determine if a language L is or not regular.

Lemma 4.1.1 (Pumping Lemma). *For every regular language L there exists a constant $n \in \mathbb{Z}^+$ (depending only on L) such that every string $w \in L$ of length at least n (n is called the "pumping length") can be written as $w = uvx$ (i.e., w can be divided into three substrings), satisfying the following conditions:*

1. *$|uv| \leq n$, where $|z|$ is the length of the string z, i.e. the number of symbols in the string z.*
2. *$v \neq \lambda$.*

3. For all $i \in \mathbb{N}$, $uv^i x \in L$.

Example 4.1.3. *Let $\Sigma = \{a, b\}$ be the alphabet, a* ***palindrome*** *is a chain w such that $w^R = w$; for example, if we consider $w = ababa$ and $v = baba$ then w is palindrome because $w = ababa = w^R$ but v does not since $v = baba \neq abab = v^R$. So, the language L of palindrome is not a regular one.*
In effect, if L was regular, there exists a pumping constant n. Let $w = a^n b a^n \in L$. Then, we can rewrite $w = uvx$ where $|v| \geq 1$, $|uv| \leq n$ and for all $i \in \mathbb{N}$, $uv^i x \in L$. Thus, u and v only have aes, i.e., there are $r \geq 0$ and $s \geq 1$ such that:

$$u = a^r,$$

$$v = a^s.$$

Then,

$$x = a^{n-(r+s)} b a^n = a^{n-r-s} b a^n.$$

By taking $i = 0$, it is concluded that $ux \in L$, but

$$ux = a^r a^{n-r-s} b a^n = a^{n-s} b a^n.$$

Because $s \geq 1$, $ux = a^{n-s} b a^n$ is not palindrome. This contradiction shows that L cannot be a regular language.

An homomorphism is a function

$$\begin{aligned} h : \Sigma^* &\rightarrow \Gamma^* \\ \lambda &\mapsto \lambda \\ a_1 a_2 \cdots a_n &\mapsto h(a_1) h(a_2) \cdots h(a_n) \end{aligned}$$

In other words, h transforms each symbol of the alphabet Σ in a string of Γ^* where Γ is another alphabet. If $h : \Sigma^* \rightarrow \Gamma^*$ is an homomorphism then, for $u, v \in \Sigma^*$ and $A, B \subseteq \Sigma^*$ we have the following:

- $h(uv) = h(u)h(v)$.
- $h(A \cup B) = h(A) \cup h(B)$.
- $h(AB) = h(A)h(B)$.
- $h(A^*) = h(A)^*$.
- Whether L is a regular language over Σ then $h(L)$ is a regular language over Γ.

A **finite automaton** is an abstract machine that processes strings in order to accept or reject them. It works by reading the symbols of a string one by one, from left to right. This device has a **control unit** which has a finite number of internal configurations called **states** (initial, halted or acceptance) and that always moves to the right. Formally a finite automaton M is defined through five parameters as follows.

Definition 4.1.1. *A* ***deterministic finite automaton (DFA)*** *M is a device characterized by a quintuple $M = (\Sigma, Q, q_0, F, \delta)$ that consists of the following elements:*

Σ, *the input alphabet. All the strings that M processes belong to Σ^*.*

Q, *a finite set of internal states (of the control unity).*

$q_0 \in Q$, *the initial state.*

$\emptyset \neq F \subseteq Q$, *the set of halting (or accepting and rejecting) states.*

$\delta : Q \times \Sigma \to Q$, *the transition function (of the machine).*

A ***non-deterministic finite automaton (NFA)*** *is a finite automaton such that its transition function is a mapping $\Delta : Q \times \Sigma \to \wp(Q)$.*

A string w is accepted if there exists a labeled path with the symbols of w such that starts in q_0 and end ends up in an acceptance state. The language accepted by M is denoted by $L(M)$. The empty word λ is accepted by a finite automaton M if and only if $q_0 \in F$ and M accepts the string $w = a_1 a_2 \cdots a_n$ if a sequence of states, $r_0, r_1, \ldots, r_n$, exists in Q with the following conditions:

1. $r_0 = q_0$.
2. For $i = 1, 2, \ldots, n-1$,

$$r_{i+1} = \begin{cases} \delta(r_i, a_{i+1}) & \text{if } M \text{ is a DFA,} \\ q \in \Delta(r_i, a_{i+1}) & \text{if } M \text{ is a NFA.} \end{cases}$$

3. $r_n \in F$.

We remark that DFA and NFA are computationally equivalent [16]. Additionally, finite automaton can be represented through a **diagram of transitions** which is a labeled digraph such that only acceptance paths are indicated and that follows the conventions of the TABLE 4.1.

Figure	**Description**
(r)	The state r. The vertices or nodes are the automaton states.
start $\longrightarrow (q_0)$	The initial state q_0.
$((q_f))$	The final state q_f.
$(q) \xrightarrow{s} (p)$	The transition $\delta(q, s) = p$.

TABLE 4.1. Conventions to represent finite automaton.

Example 4.1.4. *Given the alphabet $\Sigma = \{a, b\}$ and the language $L = \{a^{2i} b^{3j} \mid i, j \geq 0\}$. The FIGURE 4.1 shows an automaton M that accepts the language L.*

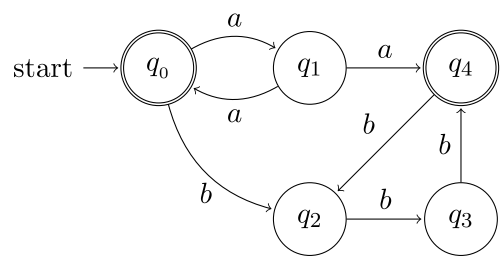

FIGURE 4.1. Automaton M that accepts the language L.

Another useful result to build an automaton that accepts two regular languages over the same alphabet is the next one.

Theorem 4.1.1. *For $i = 1, 2$, let $M_i = (\Sigma, Q_i, q_{0_i}, F_i, \delta_i)$ be DFA, then the DFA*

$$M = (\Sigma, Q_1 \times Q_2, (q_{0_1}, q_{0_2}), F_1 \times F_2, \delta)$$

where

$$\begin{aligned} \delta : Q_1 \times Q_2 \times \Sigma &\rightarrow Q_1 \times Q_2 \\ \delta(q_i, q_j, a) &\mapsto (\delta_1(q_i, a), \delta_2(q_j, a)) \end{aligned}$$

satisfies

$$L(M) = L(M_1) \cap L(M_2).$$

The word **algorithm** is a combination of the Latin word *algorismus* and the Greek word *arithmos*. Algorithms help us to process data or information in a finite number of steps which should be rigorously defined by specifying all possible circumstances that could arise and making clear (and computable) the criteria for each case. Because an algorithm is a precise list of precise steps, the order of computation and instructions are usually assumed to be listed explicitly and are described as starting "from the top" and going "down to the bottom" like in a flowchart. When there exists a decision algorithm, a problem is said to be **decidable** or **resolvable**; if not then **undecidable** or **unsolvable**.

A **generative grammar** is a finite set of production rules summarized by the quadruple $G = (V, \Sigma, S, P)$ where V is a finite alphabet whose elements are termed **variables** or **nonterminal symbols** (indicating that some production rule can be applied yet), Σ is another alphabet such that $V \cap \Sigma = \emptyset$ and whose elements are called **terminal symbols** (indicating that no production rule can be applied), S is an element of V called **start symbol**. Nonterminals are often represented by uppercase letters, terminals by lowercase letters, and the start symbol by S. P is a finite subset of $(V \cup \Sigma)^* \times (V \cup \Sigma)^*$ of **production rules** or **re-writing rules**; an element $(v, w) \in P$ is denoted $v \rightarrow w$ and means that v can be replaced by w.

The language generated by a grammar G is the set $L(G) = \{w \in \Sigma^* \mid S \stackrel{+}{\Longrightarrow} w\}$, where $u \stackrel{+}{\Longrightarrow} v$ means that v is obtained from u through one or more production rules. Grammars can be expressed by means of a derivation tree (**parse tree**); an **ambiguous grammar** is a context-free grammar for which there exists a string that can have more than one leftmost derivation or parse tree. The TABLE 4.2 summarizes each of Chomsky's types of grammars, the generated language, the type of automaton that recognizes it, and the form its rules must have.

Grammar	Languages	Automaton	Production rules
Type-0	Recursively enumerable	Turing machine	$\alpha \to \beta$ (no restrictions)
Type-1	Context-sensitive	Linear-bounded non-deterministic Turing machine	$\alpha A\beta \to \alpha\gamma\beta$
Type-2	Context-free	Non-deterministic pushdown automaton	$A \to \gamma$
Type-3	Regular	Finite state automaton	$A \to a$ and $A \to aB$

TABLE 4.2. The meaning of symbols is as follows:
a = terminal; α = terminal, non-terminal, or empty; β = terminal, non-terminal, or empty; γ = terminal or non-terminal; A = non-terminal; B = non-terminal.

For natural languages, like Spanish or English, it can be presented as phrases or sentences allowed in written and spoken communication. The variables appear bounded into angular parenthesis $\langle\rangle$, $\langle Sentence\rangle$ is the start variable of the grammar and the terminals are the words of the language. Natural languages are almost ever ambiguous because there exist many production rules and it causes that sentences can have several parse trees.

Example 4.1.5. *A grammar for Spanish language is:*

$$\begin{aligned}
\langle Sentence(S)\rangle &\to \langle Subject\rangle\langle Verb\rangle\langle DC\rangle \mid \\
&\quad\ \langle Subject\rangle\langle Verb\rangle\langle DC\rangle\langle AC\rangle \mid \\
&\quad\ \langle Subject\rangle\langle Verb\rangle\langle IC\rangle\langle AC\rangle \\
\langle Subject\rangle &\to \langle Noun\rangle \mid Juan \mid Omar \mid Ana \mid \cdots \\
\langle Direct\ Complement\ (DC)\rangle &\to \langle Prep\rangle\langle Art\rangle\langle Noun\rangle \mid \\
&\quad\ \langle Prep\rangle\langle Article\rangle\langle Noun\rangle\langle Prep\rangle\langle Art\rangle\langle Noun\rangle \mid \\
&\quad\ \langle Prep\rangle\langle Art\rangle\langle Noun\rangle\langle Prep\rangle\langle Noun\rangle\langle Adj\rangle \\
\langle Indirect\ Complement\ (IC)\rangle &\to \langle Prep\rangle\langle Art\rangle\langle Noun\rangle \mid \\
&\quad\ \langle Prep\rangle\langle Art\rangle\langle Noun\rangle\langle Adj\rangle \mid \\
&\quad\ \langle Prep\rangle\langle Noun\rangle\langle Prep\rangle\langle Noun\rangle \\
\langle Adverbial\ Complement\ (AC)\rangle &\to \langle Prep\rangle\langle Art\rangle\langle Noun\rangle \mid \\
&\quad\ \langle Prep\rangle\langle Art\rangle\langle Noun\rangle\langle Adj\rangle \mid \langle Adv\rangle \\
&\quad\ \langle Prep\rangle\langle Art\rangle\langle Noun\rangle\langle Prep\rangle\langle Art\rangle\langle Noun\rangle \\
\langle Noun\rangle &\to house \mid dog \mid book \mid pencil \mid \lambda \mid \cdots \\
\langle Adjective\ (Adj)\rangle &\to red \mid blue \mid intelligent \mid evil \mid useful \mid \lambda \mid \cdots \\
\langle Preposition\ (Prep)\rangle &\to to \mid before \mid under \mid with \mid in \mid into \mid \lambda \mid \cdots \\
\langle Article\ (Art)\rangle &\to the \mid a \mid an \mid some \mid \lambda \\
\langle Adverb\ (Adv)\rangle &\to very \mid much \mid many \mid few \mid slowly \mid fast \mid \lambda \mid \cdots \\
\langle Verb\rangle &\to write \mid wrote \mid written \mid close \mid closed \mid \lambda \mid \cdots
\end{aligned}$$

Definition 4.1.2. *A **pushdown automaton** M is an septuple $M = (Q, q_0, F, \Sigma, \Gamma, z_0, \Delta)$ that consists of the following elements:*

Q, *a finite set of internal states (of the control unity).*

q_0, *the initial state.*

$\emptyset \neq F \subseteq Q$, *the set of halting (or accepting and rejecting) states.*

Σ, *the input alphabet.*

Γ, *the working of tape alphabet. It satifies* $(\Sigma \subseteq \Gamma)$.

$z_0 \in \Gamma \setminus \Sigma$, *the initial stack symbol.*

$\delta : Q \times (\Sigma \cup \{\lambda\}) \times \Gamma \to Q \times \Gamma^*$, *the transition function (of the machine).*

*A **non-deterministic pushdown automaton (NPA)** is a pushdown automaton such that its transition function is a function $\Delta : Q \times \Sigma \to \wp_F(Q \times \Gamma^*)$, where $\wp_F(Q \times \Gamma^*)$ is the set of finite subsets of the Cartesian product $Q \times \Gamma^*$.*

A pushdown automaton reads a given input string from left to right. In each step, it chooses a transition by indexing a table by input symbol, current state, and the symbol at the top of the stack. A pushdown automaton can also manipulate the stack as part of performing a transition. The manipulation can be to push a particular symbol to the top of the stack, or to pop off the top of the stack. The automaton can alternatively ignore the stack, and leave it as it is. The FIGURE 4.2 illustrates the working of this type of automaton.

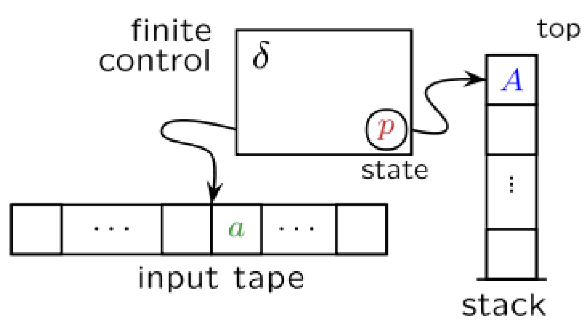

FIGURE 4.2. A diagram of a pushdown automaton.

Turing machines form the core of computability theory or recursion theory [65] and are models of automata with maximal computability ability: the control unity can move at the left or at the right and overwrite symbols in the input tape. The definition that follows is taken from De Castro's book [16].

Definition 4.1.3. *A **Turing machine** M is a device characterized by a heptuple $M = (Q, q_0, F, \Sigma, \Gamma, \sqcup, \delta)$ that consists of the following elements:*

Q, *a finite set of internal states (of the control unity).*

q_0, *the initial state.*

$\emptyset \neq F \subseteq Q$, *the set of halting (or accepting and rejecting) states.*

Σ, *the input alphabet.*

Γ, *the working of tape alphabet. It satisfies* $(\Sigma \subseteq \Gamma)$.

$\llcorner\!\lrcorner \in \Gamma \setminus \Sigma$, *the blank symbol.*

$\delta : Q \times \Gamma \to Q \times \Gamma \times D$, *the transition function (of the machine), where* $D := \{\leftarrow, \rightarrow\}$.

A Turing machine M processes input strings $w \in \Sigma^*$ and to perform such a process, the control unity of M is in the initial state q_0 scanning the first symbol of w. The Turing machine M decides a language L if for any word $w \in \Sigma^*$ either $w \in L$ and M accepts w or $w \notin L$ and M rejects w. Thus, the formal language decided by some Turing machine is

$$L(M) : \{w \in \Sigma^* \mid M \text{ accepts } w\}.$$

The transition function δ of a Turing machine M can be represented through a labeled digraph. For instance, the representation of the transition $\delta(q, a) = (p, b, \rightarrow)$ is shown in the FIGURE 4.3.

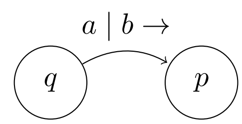

FIGURE 4.3. Transition.

In the following example (taken from [16]) a Turing machine M is built in order to accept a language and to stop up to every input is processed.

Example 4.1.6. *Let* $L = \{a^i b^i c^i \mid i \geq 0\}$ *be a language. It is a recursive language and It cannot be accepted for any pushdown automata. let* M *be a Turing machine with parameters*

$$\begin{aligned} \Sigma &= \{a, b, c\} \\ \Gamma &= \{a, b, c, X, Y, Z, \llcorner\!\lrcorner\} \\ Q &= \{q_0, q_1, q_2, q_3, q_4, q_5\} \\ F &= \{q_5\} \end{aligned}$$

and whose transition function is represented in FIGURE 4.4.

Since Turing machines have the ability of transform the input strings by erasing or overwriting symbols, they can be used as devices to compute functions [16]. Formally, a Turing machine $M = (Q, q_0, q_f, \Sigma, \Gamma, \llcorner\!\lrcorner, \delta)$ computes a function $h : \Sigma^* \to \Gamma^*$ whether for all $w \in \Sigma^*$ we obtain $q_0 w \vdash^* q_f v$ whenever $v = h(w)$. If there is a Turing machine that computes the function h, we say that h is a **Turing-computable function**. This notion can easily be extended to functions of several arguments. Specifically, a function h of k arguments is Turing-computable if for each $k-$uple $(w_1, w_2, \ldots, w_k)$ we have that $q_0 w_1 \llcorner\!\lrcorner w_2 \llcorner\!\lrcorner \ldots \llcorner\!\lrcorner w_k \llcorner\!\lrcorner \vdash^* q_f v$ whenever $v = h(w_1, w_2, \ldots, w_k)$.

Another important point of view of the Turing machines is its ability to generate languages. A Turing machine $M = (Q, q_0, F, \Sigma, \Gamma, \llcorner\!\lrcorner, \delta)$ generates the language $L \subseteq \Sigma^*$ whether:

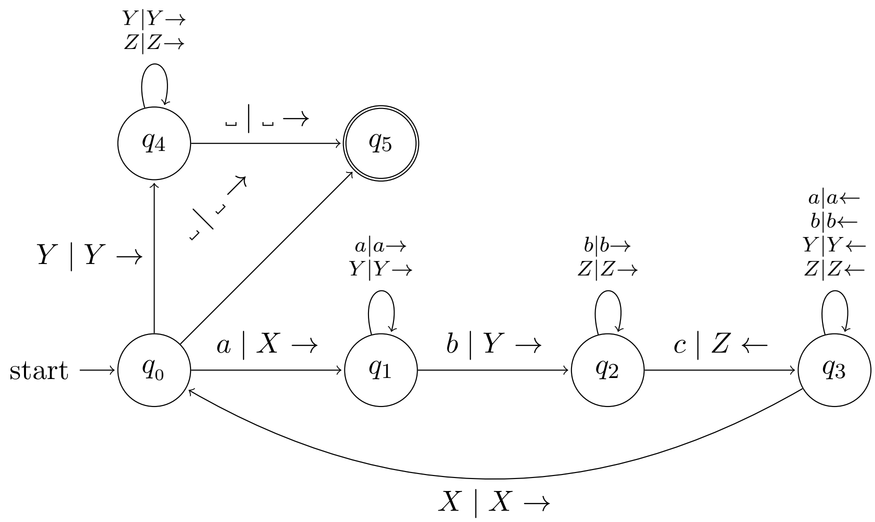

FIGURE 4.4. Turing machine.

1. M starts to operate with the blank tape in the initial state q_0.
2. Each time that M returns to its initial state q_0, there exists a string $u \in L$ written in the tape.
3. All the strings of L are generated by M.

The next example presents a Turing machine that generates all the strings of zeros and ones in lexicographic order.

Example 4.1.7. *A Turing machine M such that generates the strings of zeros and ones in lexicographic order, i.e.,* 0, 1, 00, 01, 11, 000, 001, 010, 011, 100, 101, 110, 111, 0000, 0001, … *is presented in* FIGURE *4.5.*

Note that designing a Turing machine is similar to writing a computer program since transition function of a Turing machine is just a set of instructions. The **Church-Turing thesis** says that every algorithm can be described through a Turing machine.

4.2 Extensions of the Turing Machine

In this subsection, we will briefly present three standard extensions of the Turing machine: nondeterministic and probabilistic; see [65] for more details.

Definition 4.2.1. *A **nondeterministic Turing machine** N is a Turing machine where the transition function δ is replaced by a function $\Delta : Q \times \Gamma \to \wp(Q \times \Gamma \times D)$, where $D = \{\leftarrow,\ \dagger,\ \rightarrow\}$. The accepted language by this device is*

$$L(N) = \{w \in \Sigma^* \mid N \text{ accepts } w\}.$$

Remark 25. *There are several ways to process a string and it is accepted whenever there exists a string processing which ends up in acceptance state.*

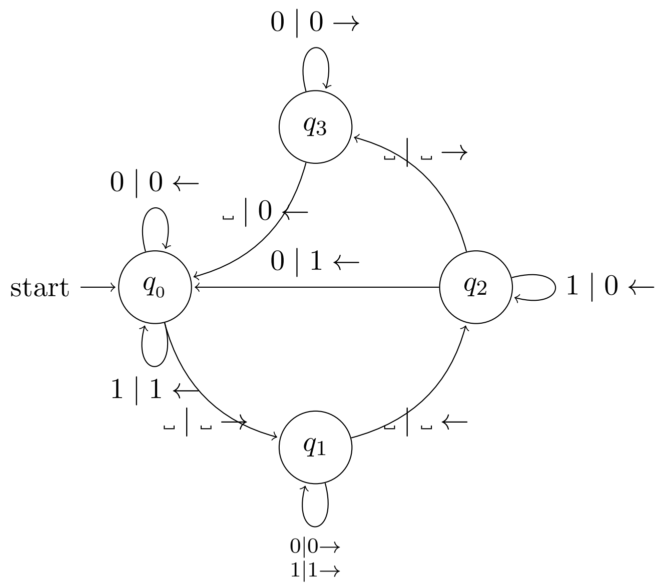

FIGURE 4.5. Turing machine M such that generates the strings of zeros and ones in lexicographic order.

Santos introduced the probabilistic Turing machine as a probabilistic extension of ordinary Turing machines [54]. Here we give the generalized definition of a probabilistic Turing machine presented by Santos in [56].

Definition 4.2.2. *A* ***probabilistic Turing machine (PTM)*** *P is a quintuple $P = (S, B, Q, p, h)$, where:*

S, *the set of symbols that the machine can print (printing alphabet);*

B, *an auxiliary alphabet;*

Q, *the set of internal states. It satisfies $S \cap Q = \emptyset$;*

$p : Q \times U \times V \times Q \to [0, 1]$, *a function that gives the probability of the next action, where $U = S \cup B$, $V = U \cup Q \cup D$, $\to \notin U \cup Q$, $\dagger \notin U \cup Q$ and $\leftarrow \notin U \cup Q$.*

$h : Q \to [0, 1]$, *a function that gives the probability that the initial state is q.*

Additionally, the functions p and h must satisfy the following conditions:

(i) For every $q \in Q$, $u \in U$, $\sum_{v \in V} \sum_{q' \in Q} p(q, u, v, q') = 1$; and

(ii) $\sum_{q \in Q} h(q) = 1$.

The function p is called the ***transition function*** *and h is termed the* ***initial distribution****.*

An expression of this machine is an element of the set $(S \cup B \cup Q)^*$. A PTM is deterministic if and only if the range of p and h is the set $\{0, 1\}$. For a PTM P, we take a similar approach likewise a deterministic Turing machine by generalizing the idea as follows.

Let $a_1, a_2, b \in S$ and $y, z \in S^*$. Then, a configuration C_1 **probabilistically yields** C_2 with probability ρ if $C_1 = yq_1a_1a_2z$, $C_2 = ybq_2a_2z$ and $p(q_1, a_1, q_2, b, \rightarrow) = \rho$ or if $C_1 = ya_2q_1a_1z$, $C_2 = yq_2a_2bz$ and $p(q_1, a_1, q_2, b, \leftarrow) = \rho$, i.e., a configuration C_1 probabilistically yields a configuration C_2 with probability ρ if the probabilistic transition function maps the two configurations to ρ. Let $T = \{T_1, T_2, \ldots, T_n\}$ be the set of all possible computational paths of P on input w and let a path T_i consist of a sequence of configurations $C_{i,1}, C_{i,2}, \ldots, C_{i,k_i}$. We say that path T_i accepts w with probability ρ_{T_i} if

1. $C_{i,1}$ is the start configuration of P on input w,
2. each $C_{i,j}$ probabilistically yields $C_{i,j+1}$ with probability ρ_j,
3. C_{i,k_i} is an accepting configuration,
4. $\rho_{T_i} = \prod_j \rho_j$.

Let $A \subseteq T$ be the set of accepting paths of P. Then, we say that PTM P accepts string w with probability $\rho(w) = \sum_{c \in A} \rho_c$.

4.3 The Fuzzy Turing machine

A number of everyday activities can be viewed as fuzzy algorithms which might contain fuzzy commands. For instance, instructions on how to treat a disease or cooking recipes. Executing fuzzy commands can be done by associating their fuzzy sets with collections of probabilities which are proportional to membership degrees and then to employ some probabilistic method to execute these commands [65].

Zadeh suggested that a fuzzy variant of Turing machine should be defined in order to characterize fuzzy algorithms since the identification of an algorithm with a Turing machine restricts the applicability of the notion of an algorithm. Since Turing machines accepts a formal language, any fuzzy Turing machine should be able to accept a fuzzy language. But, what a fuzzy language is? The following definition gives the answer.

Definition 4.3.1. *A **fuzzy language** Λ over an alphabet Σ is a fuzzy subset of Σ^*. Hence, if $w \in \Sigma^*$, then $\Lambda(w)$ is the grade of membership that indicates how much w is a member of the language.*

The following example helps to understand the previous definition.

Example 4.3.1. *Consider the set that includes all the sequences of zeros followed by one*

$$L = \{0^i 1^j \mid i \neq j \text{ and } i, j > 0\}$$

and the following function

$$\Lambda(0^i 1^j) = \begin{cases} \frac{j}{i} & i > j, \\ \frac{i}{j} & \text{otherwise.} \end{cases}$$

Both defines a fuzzy language.

Now, let Λ_1 and Λ_2 be two fuzzy languages over Σ, for all $w \in \Sigma^*$ the union, intersection and concatenation of Λ_1 and Λ_2 are defined by

$$\begin{aligned}\Lambda_1 \cup \Lambda_2(w) &:= \max\{\Lambda_1(w), \Lambda_2(w\},\\ \Lambda_1 \cap \Lambda_2(w) &:= \min\{\Lambda_1(w), \Lambda_2(w\},\\ \Lambda_1\Lambda_2(w) &:= \max_{u\in\Sigma^*}\left\{\min_{v\in\Sigma^*}\{\Lambda_1(u), \Lambda_2(v\} \mid w = uv\right\}\end{aligned}$$

respectively; see [65]. Also, if Λ is a fuzzy language in Σ, then, the **Kleene clousure** of Λ is the fuzzy subset Λ^* of Σ^* defined by

$$\Lambda^* := \max_{n\in\mathbb{N}} \Lambda^n(w),$$

where $\Lambda^n(w) := \overbrace{\Lambda\cdots\Lambda}^{n-times}(w)$.

A **fuzzy grammar** is a cuadruple $G = (\Sigma_N, \Sigma_T, P, s)$ where Σ_N and Σ_T are disjoint sets of symbols; P is a set of fuzzy production rules and $s \in \Sigma_N$ is the starting symbol. The elements of P are expressions of the form $\alpha \overset{\rho}{\to} \beta$, with $\rho > 0$ and $\alpha, \beta \in (\Sigma_N \cup \Sigma_T)^*$ and ρ is the plausibility degree that α can generate β.

Santos was the first researcher to give a formal definition of a fuzzy Turing machine; see [55]. In this text, we present the most general version of the definition o a fuzzy Turing machine: the Turing W-machine.

Definition 4.3.2. *A **Turing W-machine** M is a device characterized by an heptuple*

$$M = (\Sigma, V, \Gamma, Q, F, q_0, p),$$

where:

Σ**,** *the set of input symbols.*

V**,** *the set of output symbols.*

Γ**,** *the set of tape symbols.*

Q**,** *the set of states.*

$F \subseteq Q$**,** *is the set of terminal states.*

q_0**,** *initial state.*

$p : F \times Q \to Q \times (\Gamma \cup D)$**,** *a $W-$function where $D := \{\leftarrow,\ \dagger,\ \to\}$; neither $\leftarrow$, $\dagger$, and $\to$ belongs to Γ; and, for every $s \in \Gamma$ and $q' \neq q$, $p(q', \dagger \mid s, q) = 0$.*

In general, $p(q', z \mid c, q)$ expresses the plausibility of what happens next provided that the machine is in state q and the tape symbol c is scanned.

A $w-$**function** from A to B, where A and B are nonempty sets, is a function from $A \times B$ to an ordered semiring W; such a **ordered semiring** is a cuadruple $W = (W, \oplus, \otimes, <)$ where W is a nonempty set, $\oplus$ and $\otimes$ are binary operations on W and, $<$ is a partial ordering relation on W safisfying the following conditions:

(i) For all $x, y, z \in W$

$$\begin{aligned} x \oplus y &= y \oplus x, \\ x \otimes y &= y \otimes x, \\ x \oplus (y \oplus z) &= (x \oplus y) \oplus z, \\ x \otimes (y \otimes z) &= (x \otimes y) \otimes z, \\ x \otimes (y \oplus z) &= (x \otimes y) \oplus (x \otimes z). \end{aligned}$$

(ii) There are $0, 1 \in W$ such that for all $x \in W$

$$\begin{aligned} x \oplus 0 &= x \\ x \otimes 1 &= x. \end{aligned}$$

(iii) For all $x, y \in W$

$$\begin{aligned} x &< x \oplus y \\ y &< x \oplus y. \end{aligned}$$

Whether f is a $W-$function from A to B. Then, f is **computable** if and only if there is a Turing $W-$machine $M = (\Sigma, V, \Gamma, Q, F, q_0, p)$ such that $A = \Sigma^*$, $B = V^*$ and, for all $x \in A$ and $y \in B$, $f(y \mid x) = M(y \mid x)$. In addition, we say that f **is computable by** M or, equivalently, M **computes** f.

On the other hand, Claudio Moraga proposed a machine with fuzzy states and symbols [40].

Definition 4.3.3 (**Moraga's fuzzy Turing Machine**). *A* ***fuzzy Turing Machine*** *M is a sextuple*

$$M = (Q, \Sigma, \delta, \sqcup, q_0, F)$$

where

$\Sigma : U_\Sigma \to [0, 1]$, *is a normal fuzzy subset of symbols, where U_Σ is a universe of symbols and $\hbar \in U_\Sigma$ is a special symbol such that $\Sigma(\hbar) = 1$.*

$Q : U_Q \to [0, 1]$, *is a fuzzy subset of states, where U_Q is a finite universe of states.*

$\sqcup$ *is the blank symbol.*

q_0 *is the initial state.*

$F \subseteq Q$, *is the set of halting or accepting and rejecting states; and*

$\delta : Q \times \Sigma \to Q_0 \times \Sigma_0 \times D$ *is the transition function, where $D := \{\leftarrow, \dagger, \rightarrow\}$.*

Because of a fuzzy set $A : X \to [0, 1]$ can be charcterized by a subset of the Cartesian product $A_0 \times [0, 1]$, the function δ is alternatively expressed like:

$$\delta : (Q_0 \times [0, 1]_C) \times (\Sigma_0 \times [0, 1]_C) \to Q_0 \times \Sigma_0 \times D$$

where $[0, 1]_C \subseteq [0, 1]$ includes only the classical computble numbers [65].

We remark that by replacing δ with the relation

$$\Delta \subseteq (Q_0 \times [0,1]_C) \times (\Sigma_0 \times [0,1]_C) \times Q_0 \times \Sigma_0 \times D$$

one can obtain a nondeterministic version of the fuzzy Turing machine [65]. Unfortunately, this model is unrealistic [65] nevertheless, Moraga introduced another form of fuzzy Turing machine which is described below.

Definition 4.3.4 (**Moraga's fuzzy Turing** $W-$**Machine**). *A **fuzzy Turing** $W-$**machine** M is a sextuple*

$$M = (Q, \Sigma, \delta, \text{␣}, q_0, F)$$

where

Σ, *the input alphabet.*

Q, *a finite set of states such that* $\Sigma \cap Q = \emptyset$.

$q_0 \in Q$, *the initial state.*

$F \subseteq Q$, *is the set of halting or accepting and rejecting states,*

$\delta \subseteq (Q \times \Sigma) \times (Q \times \Sigma \times D)$ *is the transition function, where* $D := \{\leftarrow,\ \dagger,\ \rightarrow\}$*; and*

$\delta_W : (Q \times \Sigma) \times (Q \times \Sigma \times D) \rightarrow [0,1]_C$ *is a function that assigns a plausibility degree to each transition of the machine.*

The definition that follows was proposed by Wiedermann [73] and it is considered the most complete and general definition of a fuzzy Turing machine [65].

Definition 4.3.5 (**Wiedermann's Fuzzy Turing Machine**). *A **nondeterministic fuzzy Turing machine with a unidirectional tape** is a nonuple*

$$\mathfrak{F} = (Q, \Gamma, I, \Delta, \text{␣}, q_0, q_f, \mu, \wedge)$$

where,

Q *is a finite set of states.*

Γ *is a finite set of tape symbols.*

I *is a set of input symbols, i.e.,* $I \subseteq T$.

Δ *is a transition relation, i.e.,* $\Delta \subseteq Q \times \Gamma \times Q \times \Gamma \times D$, *where* $D := \{\leftarrow,\ \dagger,\ \rightarrow\}$.

␣ $\in \Gamma \setminus I$ *is the blank symbol.*

q_0 *initial state.*

q_f *final state.*

μ *is a fuzzy relation on* Δ, *that is,* $\mu : \Delta \rightarrow [0,1]$, *and*

$\wedge$ *is a T-norm.*

When γ is a fuzzy subset of Q and μ is a partial function from $Q \times \Gamma \times D$, then the resulting machine is a ***deterministic fuzzy Turing machine****.*

When a machine starts with input some string w, the characters of the string are printed on the tape starting from the leftmost cell; the scanning head is placed atop the leftmost cell, and the machine enters state q_0. If

$$S_0 \overset{\alpha_0}{\vdash} S_1 \overset{\alpha_1}{\vdash} S_2 \overset{\alpha_2}{\vdash} \cdots \overset{\alpha_{n-1}}{\vdash} S_n,$$

then S_n is reachable from S_0 in n steps. Assume that S_n is reachable from S_0 in n steps, then the plausibility degree of this **computational path** is

$$D\left((S_0, S_1, \ \ldots, S_n)\right) = \begin{cases} 1, & n = 0 \\ D\left((S_0, S_1, \ \ldots, S_{n-1} * \alpha_{n-1})\right) & n > 0. \end{cases}$$

Obviously, the value that is computed with this formula depends on the specific path that is chosen. Since the machine is nondeterministic, it is quite possible that some configuration S_n can be reached via different computational paths. Therefore, when a machine starts from S_0 and finishes at S_n in n steps, the plausibility degree of this computational path, which is called a computation, should be equal to the maximum of all possible computation paths:

$$d(S_n) = \max D\left((S_0, S_1, \ \ldots, S_n)\right).$$

In different words, the plausibility degree of the computation is equal to the plausibility degree of the computational path that is most likely to happen.

Assume that a machine starts from configuration S_0 with input the string w. Then, a computational path S_0, $S_1, \ldots,$ S_m is an accepting path of configurations if the state of S_m is q_f. In addition, the string w is accepted with degree equal to $d(S_n)$.

Definition 4.3.6. *Assume that $\mathfrak{F}$ is a fuzzy nondeterministic Turing machine. Then, an input string w is accepted with plausibility degree $e(w)$ by $\mathfrak{F}$ if and only if*

- *There is an accepting configuration from the initial configuration S_0 on input w; and*
- *$e(w) = \max_S \{d(S) \mid S$ is an accepting configuration reachable from $S_0\}$.*

Also,

Definition 4.3.7. *The fuzzy language accepted by some machine $\mathfrak{F}$ is the fuzzy set that is defined as follows:*

$$L(\mathfrak{F}) = \{(w, e(w)) \mid w \textit{ is accepted by } \mathfrak{F} \textit{ with plausibility degree } e(w)\}.$$

The class of all fuzzy languages accepted by a fuzzy Turing machine, in the sense just explained, with (classically) computable $T-$norms is denoted as Φ.

To end this chapter, we mention that a fuzzy algorithm is a collection of fuzzy instructions which are orderly executed to yield an approximate solution to a given problem. For Zadeh, an algorithm is fuzzy when its variables range over fuzzy sets, regardless of whether they are finite or uncountable infinite sets and, he suggests that a fuzzy variant of the Turing machine should be defined in order to characterize fuzzy algorithms. He noted that the state of a Turing machine at time $n+1$ depends on the state and the symbol being scanned at time n, that is,

$$q^{n+1} = f(q^n, S^n),$$

where q^n and S^n are variables ranging over the set of states and the set of tape symbols, respectively. This implies that (q^{n+1}, q^n, S^n) is an element of a subset of $Q \times Q \times \Gamma$. In the fuzzy setting, this relation should be replaced by a fuzzy one. Thus, each triple (q^{n+1}, q^n, S^n) is associated with a factual degree. In different words, there is a feasibility degree to which the machine will enter state q^{n+1} when in state q^n and the scanning device has read the symbol S^n [65]. This leads to think that every fuzzy Turing machine defines a fuzzy algorithm.

Conclusions.

- This document can be seen as an introduction to three branches of fuzzy mathematics: Fuzzy logic, Fuzzy measure theory and the theory of fuzzy computation. Several ideas from different authors were collected in order to present a self-contained work that intends to motivate readers to explore and research in this new branch of mathematics.

- Fuzzy sets theory is a very important field of investigation, as much its mathematical implications as their practical applications. Fuzzy math is not fuzzy as it is believed and it is suitable for imprecision and approximate reasoning inasmuch as it tries to formalize, for instance through the fuzzy numbers, the capability to perform a wide variety of physical and mental tasks without any measurements and any computations.

- Fuzzy sets have been successfully applied in a huge number of applications in the world since they let us deal with a wide variety of phenomena by letting us convert complex systems into simpler ones. Fuzzy models do not have the purpose of replacing stochastic methods but to provide another way of handling uncertainty because of lack of data or the use of linguistic variables.

- As a result of this work a paper, titled *Fuzzy sets: A way to represent ambiguity and subjectivity* arose and was published in Boletín de Matemáticas in 2017 [46]. In that paper the author exposes some of the basics notions of Fuzzy Mathematics, a branch of mathematics which has been in a continuous development for the last fifty years.

Future work.

- Study and develop applications of the fuzzy sets in fields like: Decision Making Problems, Fuzzy Logic, Fuzzy Control, Fuzzy Analysis, Fuzzy Algebra and so on.
- In the field of Mathematical Analysis, study and to develop methods that approximate the solution of Fuzzy Differential Equations (FDEs) like FDEs with Delay, Hybrid FDEs, Fuzzy Stochastic Differential Equations or Fuzzy Differential Inclusions. Also, the solutions of Partial Differential Equations via fuzzy methods like the adaptative ones.
- Study and develop ways to use fuzzy systems to model real-world problems.
- Develop ways of constructing automata that generate fuzzy languages.
- Study the topology of the collection of the fuzzy numbers, $\mathcal{F}_C(\mathbb{R})$.
- Study the Fuzzy integral and its applications to solve real-world problems.

Bibliography

[1] Parul Agarwal and HS Nayal, *Possibility theory versus probability theory in fuzzy measure theory*, Continuity **1** (2015), 1.

[2] Herbert Amann, Joachim Escher, and Gary Brookfield, *Analysis*, vol. 3, Springer, 2005.

[3] Gilberto Arenas-Díaz and Édgar René Ramírez-Lamus, *Fuzzy measures and fuzzy integrals*, Universitas Scientiarum **18** (2013), no. 1, 7–32.

[4] Bilal M Ayyub and George J Klir, *Uncertainty modeling and analysis in engineering and the sciences*, CRC Press, 2006.

[5] Robert G Bartle, *The elements of integration and lebesgue measure*, John Wiley & Sons, 2014.

[6] Banabas Bede, *Studies in fuzziness and soft computing 295 - Mathematics of Fuzzy Sets and Fuzzy Logic*, Springer, 2013.

[7] Barnabás Bede and Luciano Stefanini, *Generalized differentiability of fuzzy-valued functions*, Fuzzy Sets and Systems **230** (2013), 119–141.

[8] Haim Brezis, *Functional analysis, sobolev spaces and partial differential equations*, Springer Science & Business Media, 2010.

[9] James J Buckley and Esfandiar Eslami, *An introduction to fuzzy logic and fuzzy sets*, vol. 13, Springer Science & Business Media, 2002.

[10] James J Buckley and Leonard J Jowers, *Simulating continuous fuzzy systems*, Springer, 2006.

[11] Bing-Yuan Cao, *Optimal models and methods with fuzzy quantities*, Springer, 2010.

[12] Asli Celikyilmaz and I Burhan Turksen, *Modeling uncertainty with fuzzy logic*, Studies in Fuzziness and Soft Computing **240** (2009).

[13] Gustave Choquet, *Theory of capacities*, Annales de l'institut Fourier, vol. 5, Institut Fourier, 1954, pp. 131–295.

[14] Earl Cox, Michael O'Hagan, Rodman Taber, and Michael O'Hagen, *The fuzzy systems handbook with cdrom*, Academic Press, Inc., 1998.

[15] Laécio Carvalho de Barros and Rodney Carlos Bassanezi, *Tópicos de lógica fuzzy e biomatemática*, Grupo de Biomatemática, Instituto de Matemática, Estatística e Computação Científica (IMECC), Universidade Estadual de Campinas (UNICAMP), 2006.

[16] Rodrigo De Castro, *Teoría de la computación: lenguajes, autómatas, gramáticas*, Universidad Nacional de Colombia, 2004.

[17] Arthur P Dempster, *Upper and lower probabilities induced by a multivalued mapping*, The annals of mathematical statistics (1967), 325–339.

[18] D Dennenberg, *Non-additive measure and integral*, 1994.

[19] Phil Diamond, Peter E Kloeden, Peter Eris Kloeden, Australia Mathematician, and Peter Eris Kloeden, *Metric spaces of fuzzy sets: theory and applications*, World Scientific, 1994.

[20] Vladimir Dimitrov and Victor Korotkich, *Fuzzy logic: a framework for the new millennium*, vol. 81, Springer Science & Business Media, 2002.

[21] Didier Dubois, Henri M Prade, and Henri Prade, *Fundamentals of fuzzy sets*, vol. 7, Springer Science & Business Media, 2000.

[22] Didier J Dubois, *Fuzzy sets and systems: Theory and applications*, vol. 144, Academic press, 1980.

[23] Ali Emrouznejad and Marianna Marra, *Ordered weighted averaging operators 1988–2014: A citation-based literature survey*, International Journal of Intelligent Systems **29** (2014), no. 11, 994–1014.

[24] Raafat Fahmy, Hegazy Zaher, and Abd Elfattah Kandil, *A comparison between fuzzy inference systems for prediction (with application to prices of fund in egypt)*, International Journal of Computer Applications **109** (2015), no. 13.

[25] Boris A Faybishenko, *Introduction to modeling of hydrogeologic systems using fuzzy differential equations*, Fuzzy Partial Differential Equations and Relational Equations, Springer, 2004, pp. 267–284.

[26] Luis Garmendia, *The evolution of the concept of fuzzy measure*, Intelligent Data Mining, Springer, 2005, pp. 185–200.

[27] Tomasa Calvo (auth.) Gleb Beliakov, Ana Pradera, *Aggregation functions: A guide for practitioners*, Studies in Fuzziness and Soft Computing 221, Springer-Verlag Berlin Heidelberg, 2007.

[28] Michel Grabisch, *Fuzzy measures and integrals: recent developments*, Fifty Years of Fuzzy Logic and its Applications, Springer, 2015, pp. 125–151.

[29] Michael Hanss, *Applied fuzzy arithmetic*, Springer, 2005.

[30] Stefan Heinz, *Mathematical modeling*, Springer Science & Business Media, 2011.

[31] Lakhmi C Jain and NM Martin, *Fusion of neural networks, fuzzy systems and genetic algorithms: Industrial applications*, vol. 4, CRC press, 1998.

[32] Abraham Kandel, *Fuzzy mathematical techniques with applications*, (1986).

[33] George Klir and Bo Yuan, *Fuzzy sets and fuzzy logic*, vol. 4, Prentice Hall New Jersey, 1995.

[34] Rodrigo De Castro Korgi, *2 el universo latex da*, 2001.

[35] Bart Kosko, *Fuzziness vs. probability*, International Journal of General System **17** (1990), no. 2-3, 211–240.

[36] Erwin Kreyszig, *Introductory functional analysis with applications*, vol. 81, wiley New York, 1989.

[37] Kwang Hyung Lee, *First course on fuzzy theory and applications*, vol. 27, Springer Science & Business Media, 2006.

[38] Xiaodong Liu and Witold Pedrycz, *Axiomatic fuzzy set theory and its applications*, Springer, 2009.

[39] AM Mane, TD Dongale, and MS Bapat, *Application of fuzzy measure and fuzzy integral in students failure decision making*, IOSR Journal of Mathematics **10** (2014), no. 6, 47–53.

[40] Claudio Moraga, *Towards a fuzzy computability?*, Mathware & soft computing. 1999 Vol. 6 Núm. 2 [-3] (1999).

[41] John N Mordeson, *Fuzzy mathematics*, Springer, 2001.

[42] John N Mordeson, Davender S Malik, and Terry D Clark, *Application of fuzzy logic to social choice theory*, CRC Press, 2015.

[43] Juan E Moreno, Oscar Castillo, Juan R Castro, Luis G Martínez, and Patricia Melin, *Data mining for extraction of fuzzy if-then rules using mamdani and takagi-sugeno-kang fis.*, Engineering Letters **15** (2007), no. 1.

[44] James R Munkres, *Topology*, Pretince Hall, 2000.

[45] Hung T Nguyen and Michio Sugeno, *Fuzzy systems: modeling and control*, vol. 2, Springer Science & Business Media, 1998.

[46] José Rubén Niño Quevedo, *Fuzzy sets. a way to represent ambiguity and subjetivity*, Boletín de Matemáticas **24** (2017), no. 1, 57–89.

[47] Yukio Ogura, *On stochastic differential equations with fuzzy set coefficients*, Soft Methods for Handling Variability and Imprecision, Springer, 2008, pp. 263–270.

[48] Witold Pedrycz and Fernando Gomide, *Fuzzy systems engineering: toward human-centric computing*, John Wiley & Sons, 2007.

[49] Zhongfeng Qin and Xiang Li, *Fuzzy calculus for finance*, 2008.

[50] Timothy J Ross, *Fuzzy logic with engineering applications*, John Wiley & Sons, 2009.

[51] Da Ruan, *Fuzzy logic foundations and industrial applications*, Kluwer Academic Publishers, 1996.

[52] Aep Saepullah and Romi Satria Wahono, *Comparative analysis of mamdani, sugeno and tsukamoto method of fuzzy inference system for air conditioner energy saving*, Journal of Intelligent Systems **1** (2015), no. 2, 143–147.

[53] MAW Salman and NI Seno, *A comparison of mamdani and sugeno inference systems for a satellite image classification*, Anbar J. Eng. Sci (2010), 296–306.

[54] Eugene S Santos, *Probabilistic turing machines and computability*, Proceedings of the American Mathematical Society **22** (1969), no. 3, 704–710.

[55] ———, *Fuzzy algorithms*, Information and Control **17** (1970), no. 4, 326–339.

[56] ———, *Computability by probabilistic turing machines*, Transactions of the American Mathematical Society **159** (1971), 165–184.

[57] Glenn Shafer et al., *A mathematical theory of evidence*, vol. 1, Princeton university press Princeton, 1976.

[58] Stewart Shapiro, *Mathematics and reality*, Philosophy of Science (1983), 523–548.

[59] Alshalaa A Shleeg and Issmail M Ellabib, *Comparison of mamdani and sugeno fuzzy interference systems for the breast cancer risk*, International Journal of Computer, Information Science and Engineering **7** (2013), no. 10, 387–391.

[60] William Siler and James J Buckley, *Fuzzy expert systems and fuzzy reasoning*, John Wiley & Sons, 2005.

[61] Harpreet Singh, Madan M Gupta, Thomas Meitzler, Zeng-Guang Hou, Kum Kum Garg, Ashu MG Solo, and Lotfi A Zadeh, *Real-life applications of fuzzy logic*, Advances in Fuzzy Systems **2013** (2013).

[62] SN Sivanandam, Sai Sumathi, SN Deepa, et al., *Introduction to fuzzy logic using matlab*, vol. 1, Springer, 2007.

[63] Michael Spivak, *Cálculo infinitesimal*, Reverté, 1996.

[64] Michio Sugeno, *Phd thesis 1974: Theory of fuzzy integrals and its applications*, Theory of Fuzzy Integrals and Its Applications (1975).

[65] Apostolos Syropoulos, *Theory of fuzzy computation*, Springer, 2014.

[66] ———, *A (basis for a) philosophy of a theory of fuzzy computation*, arXiv preprint arXiv:1603.05162 (2016).

[67] Michio Sugeno Vicenc Torra, Yasuo Narukawa, Michio Sugeno, et al., *Non-additive measures*, Springer, 2014.

[68] Enric Trillas and Luka Eciolaza, *Fuzzy logic: an introductory course for engineering students*, vol. 320, Springer, 2015.

[69] Lefteri H Tsoukalas and Robert E Uhrig, *Fuzzy and neural approaches in engineering*, John Wiley & Sons, Inc., 1996.

[70] Édgar Rubén Muela Velasco, *Diferencias conceptuales entre la teoría de la posibilidad y los conjuntos difusos en la modelación de la incertidumbre*, Épsilon (2009), no. 13, 183–191.

[71] Xuzhu Wang, Da Ruan, and Etienne E Kerre, *Mathematics of fuzziness - basic issues*, vol. 245, Springer Science & Business Media, 2009.

[72] Zhenyuan Wang and George Klir, *Generalized measure theory*, vol. 25, Springer Science & Business Media, 2010.

[73] Jiři Wiedermann, *Characterizing the super-turing computing power and efficiency of classical fuzzy turing machines.*

[74] Mark J Wierman, *An introduction to the mathematics of uncertainty*, (2010).

[75] Ronald R Yager, *On ordered weighted averaging aggregation operators in multicriteria decisionmaking*, IEEE Transactions on systems, Man, and Cybernetics **18** (1988), no. 1, 183–190.

[76] ______, *A measure based approach to the fusion of possibilistic and probabilistic uncertainty*, Fuzzy Optimization and Decision Making **10** (2011), no. 2, 91–113.

[77] Ronald R Yager and Naif Alajlan, *Fuzzy measures in multi-criteria decision making*, Procedia Computer Science **62** (2015), 107–115.

[78] Yong Yin, Ikou Kaku, Jiafu Tang, and JianMing Zhu, *Data mining: Concepts, methods and applications in management and engineering design*, Springer Science & Business Media, 2011.

[79] Lotfi A Zadeh, *Fuzzy sets*, Information and control **8** (1965), no. 3, 338–353.

[80] ______, *A fuzzy-set-theoretic interpretation of linguistic hedges*, (1972).

[81] ______, *The concept of a linguistic variable and its application to approximate reasoning i*, Springer, 1974.

[82] ______, *The concept of a linguistic variable and its application to approximate reasoning ii*, Information sciences **8** (1975), no. 4, 301–357.

[83] ______, *The concept of a linguistic variable and its application to approximate reasoning iii*, Information sciences **9** (1975), no. 1, 43–80.

[84] ______, *Discussion: Probability theory and fuzzy logic are complementary rather than competitive*, Technometrics **37** (1995), no. 3, 271–276.

[85] ______, *Fuzzy logic = computing with words*, Fuzzy Systems, IEEE Transactions on **4** (1996), no. 2, 103–111.

[86] ______, *Fuzzy sets as a basis for a theory of possibility*, Fuzzy sets and systems **100** (1999), 9–34.

[87] ______, *Toward a generalized theory of uncertainty (gtu)—-an outline*, Information sciences **172** (2005), no. 1, 1–40.

[88] ______, *Is there a need for fuzzy logic?*, Information sciences **178** (2008), no. 13, 2751–2779.

[89] ______, *Toward extended fuzzy logic - a first step*, Fuzzy Sets and Systems **160** (2009), no. 21, 3175–3181.

[90] Lotfi A Zadeh, King-Sun Fu, and Kokichi Tanaka, *Fuzzy sets and their applications to cognitive and decision processes: Proceedings of the us–japan seminar on fuzzy sets and their applications, held at the university of california, berkeley, california, july 1-4, 1974*, Academic Press, 2014.

[91] Lotfi A Zadeh and Janusz Kacprzyk, *Computing with words in information/intelligent systems 1: Foundations*, vol. 33, Springer Science & Business Media, 1999.

[92] ______, *Computing with words in information/intelligent systems 2*, vol. 33, Springer Science & Business Media, 1999.

[93] Lotfi Asker Zadeh, *Fuzzy sets as a basis for a theory of possibility*, Fuzzy sets and systems **1** (1978), no. 1, 3–28.

[94] Lotfy A Zadeh et al., *Calculus of fuzzy restrictions*, Electronics Research Laboratory, University of California, 1975.

[95] Hassan Zarei, Ali Vahidian Kamyad, and Ali Akbar Heydari, *Fuzzy modeling and control of hiv infection*, Computational and mathematical methods in medicine **2012** (2012).

[96] Huaguang Zhang and Derong Liu, *Fuzzy modeling and fuzzy control*, Springer Science & Business Media, 2006.

[97] Hans Jürgen Zimmermann, *Fuzzy sets, decision making, and expert systems*, vol. 10, Springer Science & Business Media, 1987.

[98] ______, *Fuzzy set theory and its applications*, Springer Science & Business Media, 2001.